(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 March 2003 (27.03.2003)

PCT

(10) International Publication Number
WO 03/025215 A1

(51) International Patent Classification[7]: C12Q 1/68

(21) International Application Number: PCT/AU02/01262

(22) International Filing Date:
13 September 2002 (13.09.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PR 7685 14 September 2001 (14.09.2001) AU

(71) Applicant *(for all designated States except US)*: **THE UNIVERSITY OF QUEENSLAND** [AU/AU]; St Lucia, Queensland 4067 (AU).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **CARROLL, Bernard, John** [AU/AU]; 8/144 Ryan Street, West End, Queensland 4101 (AU). **HARRISON, Dion, Kevin** [AU/AU]; 12 Takara Court, Karana Downs, Queensland 4306 (AU). **AUNG, Hnin, Thanda** [MM/AU]; 5/64 Beatrice Street, Taringa, Queensland 4068 (AU).

(74) Agents: **HUGHES, E., John, L.** et al.; Davies Collison Cave, Level 3, 303 Coronation Drive, Milton, Queensland 4064 (AU).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: DETECTION OF DNA METHYLATION

(57) Abstract: The present invention relates generally to a method of genetic fingerprinting and more particularly to a method of establishing a methylation signature for a particular eukaryotic cell or group of cells such as in the form of tissue cells or organ-specific cells. The method of the present invention exploits the sensitivity of certain restriction enzymes to methylation. Exposure of genomic or transgene DNA to these enzymes followed by amplification results in products which establish the methylation signature or profile (*i.e.* the methylome profile) of a cell's or group of cells' genome when compared to the amplified genome in undigested form. The method of the present invention is useful *inter alia* in phenotyping a cell based on its methylation signature and provides a useful tool in functional genomics and for the design of therapeutic and trait-modifying protocols, particularly for animals and plants. The present invention can also be used to identify and map junctions between methylated and unmethylated DNA. The method of the present invention is also useful for identifying DNA methylation polymorphisms which can be used *inter alia* in diagnosis and forensics and for identifying particular genes, the function or absence of function of which are associated with a disease condition or trait. The method is also useful for monitoring the aging process of particular cells or an animal (including a human) or plant comprising such cells as well as monitoring the pluripotent or multipotent state of stem cells and development of stem cells through to maturation.

WO 03/025215 A1

Detection of DNA methylation

FIELD OF THE INVENTION

The present invention relates generally to a method of genetic fingerprinting and more particularly to a method of establishing a methylation signature for a particular eukaryotic cell or group of cells such as in the form of tissue cells or organ-specific cells. The method of the present invention exploits the sensitivity of certain restriction enzymes to methylation. Exposure of genomic or transgene DNA to these enzymes followed by amplification results in products which establish the methylation signature or profile (i.e. the methylome profile) of a cell's or group of cells' genome when compared to the amplified genome in undigested form. The method of the present invention is useful *inter alia* in phenotyping a cell based on its methylation signature and provides a useful tool in functional genomics and for the design of therapeutic and trait-modifying protocols, particularly for animals and plants. The present invention can also be used to identify and map junctions between methylated and unmethylated DNA. The method of the present invention is also useful for identifying DNA methylation polymorphisms which can be used *inter alia* in diagnosis and forensics and for identifying particular genes, the function or absence of function of which are associated with a disease condition or trait. The method is also useful for monitoring the aging process of particular cells or an animal (including a human) or plant comprising such cells as well as monitoring the pluripotent or multipotent state of stem cells and development of stem cells through to maturation.

BACKGROUND OF THE INVENTION

Bibliographic details of references provided in the subject specification are listed at the end of the specification.

Reference to any prior art in this specification is not, and should not be taken as, an acknowledgment or any form of suggestion that this prior art forms part of the common general knowledge in any country.

The Human Genome Sequencing Project has provided an invaluable and near-comprehensive database of human genes. The challenge now is to understand how the genome is expressed in a certain fashion and in particular cell types to produce the traits that ultimately determine the nature and quality of life. In recent years, it has become increasingly apparent that DNA methylation plays a key role in the regulation of gene expression in higher organisms. DNA methylation (mainly methylation of cytosines in CpG sites) is thought to direct chromatin compaction and generally the switching off of genes. The importance of DNA methylation has been highlighted by its involvement in several human diseases, including cancer, and the list of such diseases is expected to expand rapidly in the post genomics era.

Defining the genomic methylation signatures (or "methylomes") of healthy and diseased cells could lead to the understanding, early diagnosis and treatment of many human diseases.

The evolutionary roles of DNA methylation in eukaryotes have been suggested to include host defence against transposons and retrotransposons (O'Neill *et al., Nature 393:* 68-72, 1998) and developmental and tissue-specific gene regulation (Walsh and Bestor, *Genes Dev. 13:* 26-34, 1999). The role of DNA methylation in developmental gene control has been the subject of considerable controversy (Yoder *et al., Trends Genet. 13:* 335-340, 1997). A role for DNA methylation in human disease has also been widely suggested but difficult to demonstrate unequivocally. Nevertheless, there is now compelling recent evidence of its involvement in several human genetic diseases, including Rett Syndrome, ICF and Fragile X (Heinrich, *Current Biology 10:* 60-63, 2000). These diseases have been found to be linked to mutations at cytosine nucleotides, or to mutations in the genes encoding proteins that are involved in DNA methylation (Hendrich, 2000, *supra*). DNA methylation has also been implicated in carcinogenesis. Large scale genome-wide hypomethylation was detected in many tumor cells compared with their normal control cells (Laird and Jaenisch, *Annu. Rev. Genet. 30:* 441-464,1996). This deficiency in methylation is thought to cause genetic instability and subsequent tumor progression

(Lengauer *et al., Proc. Natl. Acad. Sci. USA 94:* 2545-2550, 1997). Conversely, in a number of cases, silenced tumor suppressor genes have been correlated with methylation of the promoter (Baylin and Herman, *Trends Genet. 16:* 168-174, 2000; Herman, *Semin. Cancer Biol. 9:* 359-367, 1999).

5

One of the problems limiting this important field of research has been the lack of efficient procedures for large scale assessment of DNA methylation at CpG sites around the genome.

SUMMARY OF THE INVENTION

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element or integer or group of elements or integers but not the exclusion of any other element or integer or group of elements or integers.

Nucleotide and amino acid sequences are referred to by a sequence identifier number (SEQ ID NO:). The SEQ ID NOs: correspond numerically to the sequence identifiers <400>1 (SEQ ID NO:1), <400>2 (SEQ ID NO:2), etc.

In accordance with the present invention, an amplification-based assay procedure has been developed to determine the methylation profile of nucleotides in the genome of a cell or group of cells. More particularly, the nucleotides are in the form of CpG or CpNpG sites. The ability to determine genomic and transgene methylomes in a cell or group of cells is an important tool in functional genomics and in developing the next generation of gene-expression modulating agents. Combining methylation profile with mapping enables a determination of the epigenetic consequences of internal and external stimuli. For example, methylation profiles may correlate with disease conditions or a propensity for a disease condition to develop or monitoring the aging process or the development process of cells. Furthermore, the methylation profile can be used to determine genes which either are expressed or are not expressed in certain disease states or with certain phenotypic traits. The identification of a condition or predisposition for development of a condition leads to the selection of therapeutic agents or protocols for the treatment of same.

The present invention is predicated in part on an amplification-based technology referred to herein as amplified methylation polymorphisms (AMP). The AMP technology determines the methylation profile of many thousands of CpG or CpNpG sites around the genome and provides a genetic profile of the methylation status of these sites. This genetic signature is the methylome fingerprint of a cell's or group of cells' genome. The present invention extends, however, to determination of methylation of sites other than CpG or

CpNpG sites.

The AMP technology involves amplification of DNA markers in the form of small inverted repeats comprising the CpG or CpNpG sites but where amplification depends on the methylation status of the cytosines within the amplicon or nearby.

In one particularly useful embodiment, the protocol uses a single arbitrary decamer oligonucleotide primer containing the recognition sequences of a methylation-sensitive restriction enzyme. These short oligonucleotide primers containing such recognition sequences are referred to herein as AMP primers. The recognition sequences for the methylation-sensitive restriction enzyme are located in the middle of the primer followed by up to four selective nucleotides, extending to the 3' end. AMP profiles are generated from both undigested genomic DNA and genomic DNA digested with the methylation sensitive enzyme. Comparison of the profiles from digested and undigested genomic DNA reveals three classes of AMP markers: digestion resistant (Class I) indicative of methylation, digestion sensitive (Class II) indicative of non-methylation, and digestion dependent (Class III). The nature of the last class of AMP markers is proposed to represent physically-linked *cis*-acting inhibitory sequences which suppress amplification of Class III markers from undigested template. Digestion with the enzyme removes the inhibitor from the amplicon, thereby allowing amplification. The digestion-dependent (Class III) markers are proposed to encompass a methylated restriction site or sites in the amplicon sequence flanked by a non-methylated restriction site and then the putative inhibitory sequence. Digestion-dependent markers represent, therefore, junctions between methylated and non-methylated DNA in the genome. Cloning, sequencing and mapping AMP markers shows that they often correspond to CpG islands, features known to be landmarks for genes in genomes. Southern analysis has confirmed the utility of AMP for detecting DNA methylation polymorphisms between genomes and that Class III AMPs represent the junction between methylated and unmethylated DNA in genomes.

The methylome profile of a cell's genome, therefore, is provided by the determination of AMP markers. It is proposed that changes in methylation profile indicates, determines or

causes changes in patterns of gene expression and important biological phenotypes such as cancer and aging.

The terms "methylation signature", "methylation profile" and "methylome" are used interchangeably throughout the specification and all indicate the methylation status of a cell's or group of cells' genome. They can also be used to determine the expression or absence of expression of particular genes. The method of the present invention is useful in determining methylation polymorphisms which can then be linked to particular traits and/or phenotypic characteristics or stages of aging or development and can also be useful in forensics to assist in better profiling a victim or perpetrator.

The present invention contemplates, therefore, a method for determining the methylation profile of one or more nucleotides at one or more sites within the genome of a eukaryotic cell or group of cells by obtaining a sample of genomic DNA from the cell or group of cells, digesting a sub-sample of the sample of genomic DNA with a methylation-sensitive enzyme having a recognition nucleotide sequence comprising a nucleotide capable of methylation at one or more of the above-mentioned sites, subjecting the digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a cleavable nucleotide sequence of the methylation-sensitive enzyme recognition sequence and subjecting the products of the amplification reaction to separation or other detection means relative to a control. The control comprises another sub-sample of the sample of genomic DNA not subjected to digestion by the methylation-sensitive enzyme but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products to the amplification reaction to the separation or other detection means wherein the presence of amplification products in enzyme digested and non-digested samples is indicative of a methylation-sensitive enzyme-digestion-resistant marker, the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-sensitive marker and the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-dependent marker.

The method of the present invention determines, therefore, a methylation profile comprising methylated or unmethylated CpG or CpNpGp islands characterized as being a methylation-sensitive enzyme-digestion-resistant marker, a methylation-sensitive enzyme digestion-sensitive marker or a methylation-sensitive enzyme-digestion-dependent marker. A methylation profile is developed showing the presence or absence of particular methylated CpG or CpNpGp islands as well as junction sites between methylated and unmethylated regions. These are the methylation-sensitive enzyme-digestion-dependent AMP markers.

Methylation enzymes contemplated herein include *Aat*II, *Aci*I, *Acl*I, *Age*I, *Asc*I, *Ava*I, *Bam*HI, *Bsa*A1, *Bsa*H1, *Bsi*E, *Bsi*W, *Bsr*F, *Bss*HII, *Bst*BI, *Bst*UI, *Cla*1, *Eag*I, *Hae*II, *Hga*I, *Hha*I, *Hin*PI, *Hpa*II, *Mlo*I, *Msp*I, *Nae*I, *Nar*I, *Not*I, *Nru*I and *Pml*I. *Hpa*II is particularly preferred in accordance with the present invention.

Accordingly, another aspect of the present invention provides a method for determining the methylation profile of one or more CpG or CpNpG nucleotides at one or more sites within the genome of a eukaryotic cell or group of cells by obtaining a sample of genomic DNA from the cell or group of cells, digesting a sub-sample of the sample of genomic DNA with *Hpa*II which has a recognition nucleotide sequence corresponding to or within the sites, subjecting the digested DNA to an amplification means such as polymerase chain reaction (PCR) using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a *Hpa*II cleavable nucleotide sequence and subjecting the products of the PCR to separation or other detection means relative to a control. The control comprises in this embodiment another sub-sample of the sample of genomic DNA not subjected to digestion by *Hpa*II but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products to the amplification reaction to the separation or detection means wherein the presence of PCR products in enzyme digested and non-digested samples is indicative of a *Hpa*II-digestion-resistant marker (H^r), the absence and presence of PCR products in enzyme digested and undigested samples, respectively, is indicative of a *Hpa*II-digestion-sensitive marker (H^s) and the presence and

absence of PCR products in enzyme digested and undigested samples, respectively, is indicative of a *Hpa*II-digestion-dependent marker (H^d).

Another aspect of the present invention contemplates an assay device in the form of a kit useful for determining the methylation profile of one or more nucleotides at one or more sites within the genome of a eukaryotic cell, the kit comprising in compartmental form multiple compartments each adapted to comprise one or more of buffers, diluents and enyzmes as single or multiple components which are optionally required to be admixed prior to use, the kit further comprising instructions for use, wherein the method is conducted by obtaining a sample of genomic DNA from said cell, digesting a sub-sample of said sample of genomic DNA with a methylation-sensitive enzyme having a recognition nucleotide sequence comprising a nucleotide capable of being methylated at one or more of the above-mentioned sites, subjecting the digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a cleavable nucleotide sequence of the methylation-sensitive enzyme recognition sequence and subjecting the products of the amplification reaction to separation or detection means relative to a control, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by the methylation-sensitive enzyme but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products to the amplification reaction to the separation or other detection means wherein the presence of amplification products in enzyme digested and non-digested samples is indicative of a methylation-sensitive enzyme-digestion-resistant marker, and wherein the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-sensitive marker and wherein the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-dependent marker.

Preferably, as indicated above, the methylation sensitive enzyme is *Hpa*II. Conveniently, the amplification reaction is a PCR.

The present method can also be used to detect a change in a cell's or group of cells' developmental state or a cell's or group of cells' exposure to an internal or external stimulus by detecting a change in methylation profile over time. This may be particularly relevant and useful in monitoring stem cells including stem cells in developmental transition and including embryonic stem (ES) cells, embryonic germ (EG) cells and mature or committed cells. It is also particularly useful in monitoring an animal or plant cell in tissue culture and/or following recombinant procedures. The method is also useful in forensics to identify potential characteristics of an individual based on the methylation profile of DNA samples from a victim or perpetrator. The methylation profile further provides an indicator as to the likely sensitivity of a cell or group of cells to a virus or transposon.

The method of the present invention is also useful for monitoring the state and/or extent of methylation of cloned, non-human embryos such as livestock embryos. These embryos can be checked for unwanted methylation prior to transplantation as well as after transplantation.

Alternatively, the present invention provides a method for detecting a change in a cell's or group of cells' physiological status such as from a normal state to a disease state.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a diagrammatic representation showing the AMP protocol and the Class I, II and III markers using as an example, the methylation-sensitive endonuclease restriction enzyme, *Hpa*II.

Figure 2 is a diagrammatic representation showing a Class III marker detected by the AMP protocol in one cell type but not another. Cloning, sequencing and Southern analysis has shown that inhibitory sequences are often associated with or linked to classical CpG islands.

Figure 3 is photographic representation showing different classes of AMP markers. Comparison of profiles amplified with an AMP primer from either *Hpa*II-digested (H) or undigested (U) genomic DNA. The profiles in the figure were generated from replicate DNA extractions from blood of a single human being. The figure demonstrates the reproducibility of the AMP protocol and the three classes of AMP markers: digestion sensitive (left-side arrows), digestion dependent (right-side arrows) and digestion resistant (without arrows).

Figure 4 is a photographic representation showing tissue-specific AMP profiles observed for human blood and spermatozoa. Profiles were amplified from *Hpa*II-digested (H) and undigested (U) genomic DNA of six humans (B1, B2, B3, S1, S2, and S3). Six different individuals contributed the samples; B1, B2, and B3 were blood samples, whereas S1, S2 and S3 were spermatozoa samples. Tissue specific digestion-dependent (upper panel) and digestion-sensitive (lower panel) DNA methylation polymorphisms (see arrows) were readily observed. DNA was extracted and analyzed in duplicate from each tissue sample.

Figure 5 is a diagrammatic representation showing map location of four human digestion-dependent Class III AMP markers and the flanking *Hpa*II sites. Each of these digestion-dependent markers that was cloned and sequenced was 100% homologous to sequenced regions of the human genome. The positions of the Class III amplicons (HdM 1-4) are

indicated by thick shaded boxes on the chromosomal segment. The locations of *Hpa*II sites are represented by vertical lines and the distance to, and orientation of nearby genes are indicated. Cloning, sequencing and mapping of these and additional Class III AMP markers has confirmed that they often map to CpG islands.

Figure 6 is a photographic representation showing distinct methylomes of tissues from testis, peritoneal fat pad and bone marrow of mice.

Figure 7 is a photographic representation showing widespread DNA methylation polymorphisms in mouse macrophage cell lines.

Figure 8 is representation showing that AMP is a powerful diagnostic tool for identifying stable cell lines, e.g. sub-cultures (1, 2 and 3 in replicate) of the human colon carcinoma cell line HCT116 are identical for DNA methylation and DNA sequence.

Figure 9 is a representation showing that AMP detects DNA methylation in a p53 knock-out line compared to the parent cell line HCT116. Both cell lines are stable between sub-cultures (1, 2 and 3), but AMP detects widespread DNA methylation and DNA sequence changes between the two cell lines.

Figure 10 is a photographic representation showing that Southern hybridization confirms tissue-specific AMPs between sperm (S) and blood (B).

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention provides an amplification-based assay to determine the methylome's profile in the genome of a eukaryotic cell or group of cells.

Determination of the methylation profile includes determining a methylation fingerprint or signature of global methylation within a cell's or group of cell's genome or at specific sites or regions. The terms "methylome" and "methylomes" refer to the generally multiple sites of methylation within a cell's or group of cells' genome. These sites are also generally distinctive. The methylome profile is proposed to change with developmental changes, internal and external stimuli (e.g. environmental factors, RNAi, epigenetic signals) and disease conditions (e.g. cancer). The methylome profile also directs changes in gene expression in a cell. This can lead to different developmental, physiological or metabolic states or conditions. It can also lead to the identification of one or more genes or other genetic sequences such as promoter sequences involved or associated with a disease condition or trait. This is important, for example, during proliferation, development and self-renewal of precursor cells such as embryonic stem (ES) cells, embryonic germ (EG) cells and post-natal (e.g. adult) or committed cells. It is also relevant in monitoring plant or animal cells in tissue culture and/or following recombinant DNA procedures as well as monitoring the extent of methylation of animal embryos such as in the livestock industry. The embryos may be checked at pre- and post-transplantation stages.

Accordingly, one aspect of the present invention contemplates a method for determining the methylation profile of one or more nucleotides at one or more sites within the genome of a eukaryotic cell or group of cells, said method comprising obtaining a sample of genomic DNA from said cell or group of cells, digesting a sub-sample of said sample of genomic DNA with a methylation-sensitive enzyme having a recognition nucleotide sequence comprising a nucleotide capable of methylation at one or more of the above-mentioned sites, subjecting said digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a cleavable nucleotide sequence of the methylation-sensitive enzyme recognition sequence

and subjecting the products of the amplification reaction to separation or other detection means relative to a control but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of the amplification reaction to the separation and/or detection means, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by the methylation-sensitive enzyme wherein the presence of amplification products in enzyme digested and non-digested samples is indicative of a methylation-sensitive enzyme-digestion-resistant marker, the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-sensitive marker and the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-dependent marker.

The method of this aspect of the present invention involves obtaining a sample of DNA and digesting one sub-sample with a methylation-sensitive enzyme and keeping another sub-sample undigested. Both samples are then subjected to amplification with a primer having a complementary sequence to a recognition sequence for the methylation-sensitive enzyme. If methylation has occurred at a particular site, the complementary sequence of the primer is intact as the enzyme is unable to digest the nucleotide sequence at that site. An amplification product is, therefore, present in the digested and undigested samples. If methylation has not occurred, digestion with the methylation sensitive enzyme destroys the primer-binding site. Consequently, the undigested sample will have an amplification product but the digested sample will not. Finally, if the methylated site is flanked or proximal to an amplification inhibiting region, the digestion will remove that region to enable an amplification product but this will be absent in the undigested sample due to the inhibition. This is regarded as a junction between methylated and unmethylated DNA. Thus, the AMP protocol detects junctions between methylated and unmethylated DNA in the genome.

In a preferred aspect of the present invention, a global methylation signature is determined using methylation-sensitive restriction endonuclease sites within islands of repeating

nucleotides. The most preferred islands are referred to as CpG or CpNpG islands where N is any nucleotide. The most preferred methylation-sensitive restriction endonuclease sites within or comprising CpG or CpNpG islands are *Hpa*II sites. A *Hpa*II site is CCGG.

Accordingly, another aspect of the present invention contemplates a method for determining the methylation profile of one or more CpG or CpNpG nucleotides at one or more sites within the genome of a eukaryotic cell, said method comprising obtaining a sample of genomic DNA from said cell, digesting a sub-sample of said sample of genomic DNA with *Hpa*II which has a recognition nucleotide sequence corresponding to or within said sites, subjecting said digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a *Hpa*II cleavable nucleotide sequence and subjecting the products of the amplification reaction to separation or other detection means relative to a control but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of the amplification reaction to the separation or other detection means, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by *Hpa*II, wherein the presence of amplification products in enzyme digested and non-digested samples is indicative of a *Hpa*II-digestion-resistant marker (H^r), the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a *Hpa*II-digestion-sensitive marker (H^s) and the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a *Hpa*II-digestion-dependent marker (H^d).

Whilst *Hpa*II, and its functional equivalents, are the most convenient methylation-sensitive enzyme for use in the present invention up to the present time, one skilled in the art would immediately recognize that other enzymes may be employed with recognition nucleotide sequences within, comprising or adjacent to a CpG or CpNpG nucleotide sequences. A recognition nucleotide sequence is the sequence of nucleotides which is essential for the cleavage by a restriction endonuclease. A recognition nucleotide sequence is a sequence which is cleaved by the enzyme. A list of other enzymes which may be used include *Aat*II,

*Aci*I, *Acl*I, *Age*I, *Asc*I, *Ava*I, *Bam*HI, *Bsa*A1, *Bsa*H1, *Bsi*E, *Bsi*W, *Bsr*F, *Bss*HII, *Bst*BI, *Bst*UI, *Cla*1, *Eag*I, *Hae*II, *Hga*I, *Hha*I, *Hin*PI, *Mlo*I, *Msp*I, *Nae*I, *Nar*I, *Not*I, *Nru*I and *Pml*I.

The term "digestion" means that the DNA and restriction enzyme are brought together in physical biochemical contact for a time and under conditions sufficient for the cleavable nucleotide sequences to undergo cleavage. Digestion is generally for a time to permit cleavage of substantially all cleavable sites specific for the particular enzyme.

The term "genomic DNA" includes all DNA in a cell, group of cells, or in an organelle of a cell and includes exogenous DNA such a transgenes introduced into a cell.

Any amplification methodology may be employed in accordance with the present invention. Amplification methodologies contemplated herein include the polymerase chain reaction (PCR) such as disclosed in U.S. Patent Nos. 4,683,202 and 4,683,195; the ligase chain reaction (LCR) such as disclosed in European Patent Application No. EP-A-320 308 and gap filling LCR (GLCR) or variations thereof such as disclosed in International Patent Publication No. WO 90/01069, European Patent Application EP-A-439 182, British Patent No. GB 2,225,112A and International Patent Publication No. WO 93/00447. Other amplification techniques include Qβ replicase such as described in the literature; Stand Displacment Amplification (SDA) such as described in European Patent Application Nos. EP-A-497 272 and EP-A-500 224; Self-Sustained Sequence Replication (3SR) such as described in Fahy *et al.* (*PCR Methods Appl. 1(1):* 25-33, 1991) and Nucleic Acid Sequence-Based Amplification (NASBA) such as described in the literature.

Some amplification reactions, for example, PCR and LCR, involve cycles of alternately high and low set temperatures, a process known as "thermal cycling". PCR or "polymerase chain reaction" is an amplification reaction in which a polymerase enzyme, usually thermostable, generates multiple copies of the original sequence by extension of a primer using the original nucleic acid as a template. PCR is described in more detail in U.S. Patent Nos. 4,683,202 and 4,683,195. LCR or "ligase chain reaction" is a nucleic acid

amplification reaction in which a ligase enzyme, usually thermostable, generates multiple copies of the original sequence by ligating two or more oligonucleotide probes while they are hybridized to the target. LCR and its variation, Gap LCR, are described in more detail in European Patent Application Nos. EP-A-320-308 and EP-A-439-182 and International Patent Publication No. WO 90/100447 and elsewhere.

The PCR amplification process is the most preferred in practising the present invention.

Accordingly, in a preferred aspect of the present invention, there is provided a method for determining the methylation profile of one or more CpG or CpNpG nucleotides at one or more sites within the genome of a eukaryotic cell or group of cells, said method comprising obtaining a sample of genomic DNA from said cell or group of cells, digesting a sub-sample of said sample of genomic DNA with *Hpa*II which has a recognition nucleotide sequence corresponding to or within said sites, subjecting said digested DNA to a polymerase chain reaction (PCR) using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a *Hpa*II cleavable nucleotide sequence and subjecting the products of the PCR to separation or other detection means relative to a control but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of the amplification reaction to the separation or other detection means, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by *Hpa*II wherein the presence of PCR products in enzyme digested and non-digested samples is indicative of a *Hpa*II-digestion-resistant marker (H^r), the absence and presence of PCR products in enzyme digested and undigested samples, respectively, is indicative of a *Hpa*II-digestion-sensitive marker (H^s) and the presence and absence of PCR products in enzyme digested and undigested samples, respectively, is indicative of a *Hpa*II-digestion-dependent marker (H^d).

The term "separation means" extends to any profiling protocol aimed at distinguishing between different amplification products. The term "separation means" may also be read as "separation or other detection means". Gel electrophoresis is the most convenient as this permits comparison of individual bands following separation. The term "gel" includes

agarose and polyacrylamide amongst others. A range of other separation means may also be applied such as those involving spectrophotometric, fluorometric and mass spectrographic methods. These latter methods are particularly useful for detecting a specific amplicon or amplicons.

In one particular embodiment, the presence or absence of AMP markers is assayed in real-time using a real-time PCR. As indicated above, PCR is well known to those in the art. A real-time PCR reaction includes a PCR reaction in which the products are detected during the PCR reaction. Preferably, the detection of the PCR product or products is made during the linear phase of a PCR reaction. It is known to those skilled in the art that, during the linear phase of a PCR, the quantity of a product or amplicon formed is proportional to the quantity of target sequence present at the initiation of the PCR.

In a preferred real-time PCR, a labeled probe molecule capable of hybridizing to the target sequence (e.g. an AMP marker) is used in the PCR. In subsequent amplification rounds, the probe is displaced from the amplicon by the 5' exonuclease activity of Taq polymerase. Generally, but not exclusively, the probes are labeled at the 5' nucleotide and the exonuclease activity releases the 5' probe molecule. Release of the labeled nucleotide from the probe results in the ability to detect the presence of a specific PCR product. Real-time PCR assays are well known to those skilled in the art and include those disclosed in U.S. Patent No. 6,277,638.

The term "label" means a moiety which is conjugated to a probe or an oligonucleotide moiety. A labeled probe or oligonucleotide moeity can be detected, identified or captured by virtue of the label of moiety which is conjugated to the probe. A "directly detectable" label is a signal-producing label which is capable of detection either directly or through its interaction with a substance such as a substrate (in the case of an enzyme), a light source (in the case of a fluorescent compound) or a photomultiplier tube (in the case of a radioactive or chemiluminescent compound). A "proximity label" is one of at least two labels which interact with each other to produce a detectable signal when the proximity labels are brought together. Typically, a first proximity label is used in combination with a

corresponding second proximity label. An "indirectly detectable" label is a substance which in and of itself does not provide a signal but which can be utilized to identify an oligonucleotide to which the indirectly detectable label is attached. For example, biotin can be an indirectly detectable label.

DNA molecules are said to have a "5'-terminus" (5' end) and a "3'-terminus" (3' end) because nucleic acid phosphodiester linkages occur to the 5' carbon and 3' carbon of the pentose ring of the substituent mononucleotides. A terminal nucleotide as used herein is the nucleotide at the end position of the 3'- or 5'- terminus. As used herein, a nucleotide sequence, even if internal to a larger oligonucleotide or polynucleotide, can also be said to have 5'- and 3'- ends.

Reference herein to a "multiplex PCR" refers to a real-time PCR that contains more than one probe. In this embodiment, AMP marker separation is achieved by monitoring either the accumulation of a specific labeled nucleotide or monitoring the depletion of a specifically labeled nucleotide from the PCR.

Accordingly, another aspect of the present invention contemplates a method for determining the methylation profile of one or more nucleotides at one or more sites within the genome of a eukaryotic cell or group of cells, said method comprising obtaining a sample of genomic DNA from said cell or group of cells, digesting a sub-sample of said sample of genomic DNA with a methylation-sensitive enzyme having a recognition nucleotide sequence comprising a nucleotide capable of methylation at one or more of the above-mentioned sites, subjecting said digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a cleavable nucleotide sequence of the methylation-sensitive enzyme recognition sequence and subjecting the products of the amplification reaction to real-time PCR to distinguish amplification products relative to a control but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of the amplification reaction to the separation or other detection means, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by the

methylation-sensitive enzyme wherein the presence of amplification products in enzyme digested and non-digested samples is indicative of a methylation-sensitive enzyme-digestion-resistant marker, the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-sensitive marker and the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-dependent marker.

More particularly, the present invention provides a method for determining the methylation profile of one or more CpG or CpNpG nucleotides at one or more sites within the genome of a eukaryotic cell or group of cells, said method comprising obtaining a sample of genomic DNA from said cell, digesting a sub-sample of said sample of genomic DNA with *Hpa*II which has a recognition nucleotide sequence corresponding to or within said sites, subjecting said digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a *Hpa*II cleavable nucleotide sequence and subjecting the products of the amplification reaction to distinguish amplification products relative to a control but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of the amplification reaction to the separation or other detection means, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by *Hpa*II, wherein the presence of amplification products in enzyme digested and non-digested samples is indicative of a *Hpa*II-digestion-resistant marker (H^r), the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a *Hpa*II-digestion-sensitive marker (H^s) and the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a *Hpa*II-digestion-dependent marker (H^d).

Reference herein to a "probe molecule" refers to a labeled nucleotide molecule which is capable of hybridizing to a target sequence. A target sequence is generally a DNA amplicon produced by a first amplification reaction following restriction digestion by a methylation-sensitive enzyme. Methods for the production of nucleic acid probes suitable

for use in PCR are well known in the art. Methods for the determination of a suitable size and sequence of an oligonucleotide are also well known in the art.

In one particularly useful embodiment, the probe molecule is fluorescently labeled at its 5' end.

Fluorescent nucleotide analogs are widely known and commercially available from several sources. An exemplary source is NEN (trademark) Life Science Products (Boston, Massachusetts, USA) which offers dideoxy-, deoxy- and ribonucloetide analogs labeled with fluorescein, courmarin, tetramethylrhodamine, naphthofluorescein, pyrene, Texas Red (registered trademark) and Lissamine (trademark). Other suppliers include Amersham Pharmacia Biotech (Uppsala, Sweden; Piscataway, NJ, USA) and MBI Fermentas, Inc. (Amherst, NY, USA).

An advantage to using fluorescent labels and fluorescence spectroscopic analysis is that there are multiple different labels available. Such different labels can be particularly useful in a multiplex embodiment of the present invention. Different fluorescent labels may be used in different probes so that the detection of a particular fluorescently-labeled nucleotide can be distinguished from others and used to deduce which AMP markers are present.

For example, fluorescein has a 488 nm excitation and 520 nm emission wavelength, whereas rhodamine (in the form of tetramethylrhodamine) has 550 nm excitation and 575 nm emission wavelength. A fluorescence detector provides an excitation source and an emission detector. The emission wavelengths of 520 nm and 575 nm are easily distinguishable using fluorescence spectroscopy.

On a per molecule basis, fluorescence spectroscopy is about 10-fold more sensitive than absorbance spectroscopy. A very wide variety of fluorescence spectroscopy-based detectors are commercially available for reading fluorescence values of single tubes, flow cells and multi-well plates, amongst others. For example, Labsystems Multiskan models of

microplate readers are widely available with a spectral range of 400-750 nm and filters for 340, 405, 414, 450, 492, 540, 620 and 690 nm (e.g. Fisher Scientific, Pittsburgh, PA, USA).

In a related embodiment, the present invention contemplates the use of fluorescent resonant energy transfer (FRET) for the detection of a PCR amplicon. FRET is a form of molecular energy transfer by which energy is passed between a donor molecule and an acceptor molecule. FRET arises from the properties of certain chemical compounds. When excited by exposure to particular wavelengths of light, they emit light (i.e. they fluoresce) at a different wavelength. Such compounds are termed "fluorophores". In FRET, energy is passed over a long distance between a donor molecule which is a fluorophore and an acceptor molecule. The donor absorbs a photon and transfers energy to the acceptor.

Molecules that are commonly used in FRET include fluorescein, 5-carboxyfluorescein (FAM), 2'7' dimethoxy-4'5'-dichloro-6-carbosyfjuorescein (JOI), rhodamine, 6-carboxy-rhodamine (R6G), N,N,N,N-tetramethyl-6-carboxyrhodamine (TAMRA), 6-carboxy-X-rhodamine (ROX), 4-(4'-dimethylaminophenylazo) benzoic acid (DABCYL) and 5-(2'-aminoethyl)aminonaphthalene-1-sulfonic acid (EDANS).

In another aspect of the present invention, absorbance spectrographic analysis is contemplated to provide an analysis of AMP markers generated in a PCR reaction thereby providing for the determination of the methylation status of a gene. In this aspect of the present invention, a chromatographic separation with light absorbance detection of an AMP marker amplicon in a PCR mixture can be used to measure the quantity of PCR product. In this case, the second PCR employs a specific probe/primer to identify a particular amplicon. In an illustrative embodiment, a multiplexed assay for the presence of several different nucleic acid target sequences in a sample is analyzed by absorbance spectroscopy. Several labeled probes to various nucleic acid target sequences are added to a nucleic acid sample. The labels on the probes may be various nucleotide analogs, a different one for each probe. A PCR reaction is conducted and the quantity of a PCR product is measured by adsorbance spectroscopy. In this aspect of the present invention,

the PCR is directly loaded onto a pre-equilibrated high pressure liquid chromatography (HPLC) column and eluted under conditions that separate the nucleotide analogs from the natural nucleotides. Useful media for chromatographic separation of nucleotides, bases and nucleosides include reverse phase media (such as a reverse phase C18 column or ODS-80T_m or ODS-120T TSK-GEL by TosoHaas [Montgomeryville, PA, USA]), anion exchange media (such as DEAE-25SW or SP-25W TSK-GEL by TosoHaas [Montgomeryville, PA, USA]) or affinity media (such as Boronate-5PW TSK-GEL by [Montgomeryville, PA, USA]).

The HPLC column is fitted with an absorbance detector to monitor the column effluent, hence, "absorbance spectroscopy" for this type of analysis. Typical wavelengths for monitoring HPLC detection of nucleotides are 250 nm, 260 nm and 280 nm. Thus, the quantity of a PCR product may be determined in real-time by periodic sampling of the PCR.

Identification of the AMP marker PCR products can be accomplished by comparison of the retention times (as monitored by absorbance of effluent at various times) of standards of the nucleotide analogs separated on the same HPLC column under the same conditions. Alternatively, the identity of the nucleotide analogs collected in separate fractions after chromatographic separation (as determined by monitoring the absorbance of the column effluent) can be determined by other standard analytical methods such as gel electrophoresis. In an alternative aspect, the PCR mixture is separated on a gas chromatograph fitted with an absorbance detector to monitor column effluent.

In another embodiment, the AMP marker or multiplex of AMP markers generated in a single PCR reaction are hybridized to a microarray of potentially complementary nucleotide sequences such as DNA sequences. As a single oligonucleotide of the present invention (comprising, for example, a *Hpa*II site) may generate a plurality of AMP markers belonging to different AMP classes, a rapid method is required to determine the nucleotide sequence on the AMP markers. The nucleotide sequence identity of a particular AMP marker can be determined by reference, for example, to a known genomic DNA

sequence. Suitable arrays of DNA for the analysis of AMP markers include microarrays of human genes and/or contigs derived from genomic libraries, or particularly, arrays or CpG islands in genomes. Hybridization of a PCR product comprising one or a plurality of AMP markers facilitates the isolation and identification of genetic sequences.

A "nucleic acid" as used herein, is a covalently linked sequence of nucleotides in which the 3' position of the phosphorylated pentose of one nucleotide is joined by a phosphodiester group to the 5' position of the pentose of the next nucleotide and in which the nucleotide residues are linked in specific sequence; i.e. a linear order of nucleotides. A "polynucleotide" as used herein, is a nucleic acid containing a sequence that is greater than about 100 nucleotides in length. An "oligonucleotide" as used herein, is a short polynucleotide or a portion of a polynucleotide. An oligonucleotide typically contains a sequence of about two to about one hundred bases. The word "oligo" is sometimes used in place of the word "oligonucleotide". The term "oligo" also includes a particularly preferred primer length in the practice of the present invention of up to about 10 nucleotides.

As used herein, the term "primer" refers to an oligonucleotide or polynucleotide that is capable of hybridizing to another nucleic acid of interest under particular stringency conditions. A primer may occur naturally as in a purified restriction digest or be produced synthetically, by recombinant means or by PCR amplification. The terms "probe" and "primers" may be used interchangeably, although to the extent that an oligonucleotide is used in a PCR or other amplification reaction, the term is generally "primer". The ability to hybridize is dependent in part on the degree of complementarity between the nucleotide sequence of the primer and complementary sequence on the target DNA.

The terms "complementary" or "complementarity" are used in reference to nucleic acids (i.e. a sequence of nucleotides) related by the well-known base-pairing rules that A pairs with T or U and C pairs with G. For example, the sequence 5'-A-G-T-3' is complementary to the sequence 3'-T-C-A-5' in DNA and 3'-U-C-A-5' in RNA. Complementarity can be "partial" in which only some of the nucleotide bases are matched according to the base pairing rules. On the other hand, there may be "complete" or "total" complementarity

between the nucleic acid strands when all of the bases are matched according to base-pairing rules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands as known well in the art. This is of particular importance in detection methods that depend upon binding between nucleic acids, such as those of the invention. The term "substantially complementary" is used to describe any primer that can hybridize to either or both strands of the target nucleic acid sequence under conditions of low stringency as described below or, preferably, in polymerase reaction buffer (Promega, M195A) heated to 95°C and then cooled to room temperature. As used herein, when the primer is referred to as partially or totally complementary to the target nucleic acid, that refers to the 3'-terminal region of the probe (i.e. within about 10 nucleotides of the 3'-terminal nucleotide position).

Reference herein to a stringency in relation to hybridization includes and encompasses from at least about 0 to at least about 15% v/v formamide and from at least about 1 M to at least about 2 M salt for hybridization, and at least about 1 M to at least about 2 M salt for washing conditions. Generally, low stringency is at from about 25-30°C to about 42°C. The temperature may be altered and higher temperatures used to replace formamide and/or to give alternative stringency conditions. Alternative stringency conditions may be applied where necessary, such as medium stringency, which includes and encompasses from at least about 16% v/v to at least about 30% v/v formamide and from at least about 0.5 M to at least about 0.9 M salt for hybridization, and at least about 0.5 M to at least about 0.9 M salt for washing conditions, or high stringency, which includes and encompasses from at least about 31% v/v to at least about 50% v/v formamide and from at least about 0.01 M to at least about 0.15 M salt for hybridization, and at least about 0.01 M to at least about 0.15 M salt for washing conditions. In general, washing is carried out $T_m = 69.3 + 0.41$ (G+C)% (Marmur and Doty, *J. Mol. Biol. 5:* 109,1962). However, the T_m of a duplex DNA decreases by 1°C with every increase of 1% in the number of mismatch base pairs (Bonner and Laskey, *Eur. J. Biochem. 46:* 83, 1974). Formamide is optional in these hybridization conditions. Accordingly, particularly preferred levels of stringency are defined as follows: low stringency is 6 x SSC buffer, 0.1% w/v SDS at 25-42°C; a moderate stringency is 2 x

SSC buffer, 0.1% w/v SDS at a temperature in the range 20°C to 65°C; high stringency is 0.1 x SSC buffer, 0.1% w/v SDS at a temperature of at least 65°C.

The present invention provides, therefore, a methylation profile of the genome of a eukaryotic cell or group of cells, the methylation profile comprising the presence or absence of methylation at particular sites and/or junctions between methylated and unmethylated regions, said method comprising obtaining a sample of genomic DNA from said cell or group of cells, digesting a sub-sample of said sample of genomic DNA with a methylation-sensitive enzyme having a recognition nucleotide sequence comprising a nucleotide capable of methylation at one or more of the above-mentioned sites, subjecting said digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a cleavable nucleotide sequence of the methylation-sensitive enzyme recognition sequence and subjecting the products of the amplification reaction to separation or other detection means relative to a control but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of the amplification reaction to the separation or other detection means, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by the methylation-sensitive enzyme wherein the presence of amplification products in enzyme digested and non-digested samples is indicative of a methylation-sensitive enzyme-digestion-resistant marker, the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-sensitive marker and the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-dependent marker.

The present invention also contemplates kits for determining the methylation profile of one or more nucleotides at one or more sites within the genome of a eukaryotic cell or group of cells. The kits may comprise many different forms but in a preferred embodiment, the kits are designed to create products capable of resolution *via* electrophoresis.

The kit may also comprise instructions for use.

Accordingly, another aspect of the present invention contemplates an assay device in the form of a kit useful for determining the methylation profile of one or more nucleotides at one or more sites within the genome of a eukaryotic cell or group of cells, said kit comprising in compartmental form multiple compartments each adapted to comprise one or more of buffers, diluents and enyzmes as single or multiple components which are optionally required to be admixed prior to use, said kit further comprising instructions for use, wherein the method is conducted by obtaining a sample of genomic DNA from said cell or group of cells, digesting a sub-sample of said sample of genomic DNA with a methylation-sensitive enzyme having a recognition nucleotide sequence comprising a nucleotide capable of being methylated at one or more of the above-mentioned sites, subjecting said digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a cleavable nucleotide sequence of the methylation-sensitive enzyme recognition sequence and subjecting the products of the amplification reaction to separation or other detection means relative to a control but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of the amplification reaction to the separation or other detection means, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by the methylation-sensitive enzyme, wherein the presence of amplification products in enzyme digested and non-digested samples is indicative of a methylation-sensitive enzyme-digestion-resistant marker, and wherein the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-sensitive marker and wherein the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-dependent marker.

More particularly, another aspect of the present invention contemplates an assay device in the form of a kit useful for determining the methylation profile of one or more nucleotides at one or more sites within the genome of a eukaryotic cell or group of cells, said kit

comprising in compartmental form multiple compartments each adapted to comprise one or more of buffers, diluents and enyzmes as single or multiple components which are optionally required to be admixed prior to use, said kit further comprising instructions for use, wherein the method is conducted by obtaining a sample of genomic DNA from said cell or group of cells, digesting a sub-sample of said sample of genomic DNA with *Hpa*II which has a recognition nucleotide sequence corresponding to or within said sites, subjecting said digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a *Hpa*II cleavable nucleotide sequence and subjecting the products of the amplification reaction to separation or other detection means relative to a control but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of the amplification reaction to the separation or other detection means, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by *Hpa*II, wherein the presence of amplification products in enzyme digested and non-digested samples is indicative of a *Hpa*II-digestion-resistant marker (H^r), the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a *Hpa*II-digestion-sensitive marker (H^s) and the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a *Hpa*II-digestion dependent marker (H^d).

Conveniently, the kits are adapted to contain compartments for two or more of the above-listed components. Furthermore, buffers, nucleotides and/or enzymes may be combined into a single compartment.

The Taq polymerase is an example of a suitable DNA polymerase which is thermostable and which is used in PCR and other amplification reactions. The thermostable DNA polymerase is used in an amount sufficient for a hybridized probe to release an identifier nucleotide. This amount may vary with the enzyme used and also with the temperature at which depolymerization is carried out. An enzyme of a kit is typically present in an amount sufficient to permit the use of about 0.1 to 100 U/reaction; in particularly preferred

embodiments, the concentration is about 0.5 U/reaction. The Stoffel fragment is another example of a suitable DNA polymerase and is in fact preferred over Taq for AMPs.

As stated above, instructions optionally present in such kits instruct the user on how to use the components of the kit to perform the various methods of the present invention. It is contemplated that these instructions include a description of the detection methods of the subject invention, including detection by gel electrophoresis.

The present invention further contemplates kits which contain a primer for a nucleic acid target of interest with the primer being complementary to a predetermined nucleic acid target, such as comprising the complementary sequence or sense sequence of a *Hpa*II recognition sequence. In another embodiment, the kit contains multiple primers or probes, each of which contains a different base at an interrogation position or which is designed to interrogate different target DNA sequences. In a contemplated embodiment, multiple probes are provided for a set of nucleic acid target sequences that give rise to analytical results which are distinguishable for the various probes.

It is contemplated that a kit comprises a vessel containing a purified and isolated enzyme whose activity is to release one or more nucleotides from the 3' terminus of a hybridized nucleic acid probe and a vessel containing pyrophosphate. In one embodiment, these items are combined in a single vessel. It is contemplated that the enzyme is either in solution or provided as a solid (e.g. as a lyophilized powder); the same is true for the pyrophosphate. Preferably, the enzyme is provided in solution. Some contemplated kits contain labeled nucleic acid probes. Other contemplated kits further comprise vessels containing labels and vessels containing reagents for attaching the labels. Microtiter trays are particularly useful and these may comprise from two to 100,000 wells or from about six to about 10,000 wells or from about six to about 1,000 wells.

The methylation profile of a cell's or group of cells' genome may be in any location including promoter regions, 5' regions prior to the authentic transcription start site, within coding (i.e. exon) and non-coding (i.e. intron) regions of genes and in the 3' regions

following the authentic termination signal. The methylome may also be in regulatory genes, structural genes and in non-protein-encoding genes.

The present invention has a range of utilities. For example, a number of cancers have been detected in which previously silenced genes (such as silenced at birth) are now expressed. A change in methylation profile, especially if mapped to cancer-associated genes, provides a diagnostic application for the AMP technology. Indeed, many complex medical disorders are likely to have a methylation component. For example, during hypertrophy in heart disease, there is a reappearance of expression of juvenile-expressed genes. This is presumably due to changes in methylation profile. In addition, methylation profiles may be implicated in the aging process as well as in cell development such as the development of stem cells (e.g. ES cells, post-natal stem cells) or EG cells to mature cells or cells of a developmental state in between ES cells and mature cells. Still further, the extent or type of methylation may be predictive of the sensitivity or otherwise of a cell's or group of cells' sensitivity to viruses or to transposon mutagenesis or transgene expression.

In a related embodiment, changes in methylation profile may, in one embodiment, be associated with RNAi. It is proposed that RNAi is induced by double-stranded RNA whether naturally occurring or artificially produced *via* co-suppression constructs. A change in methylome profile may be a result of a change in RNAi-specific modulation of gene expression. The detection of a change in methylome profile may provide a target for gene therapy. Furthermore, a change in methylation profile may be directly or indirectly induced.

As indicated above, methylation may also be a key element in the aging process. Modification of methylation profiles may delay the onset of or ameliorate the effects of age-related conditions and illnesses. Support for this proposition is in the change in methylation profiles in identical twins. The detection of age-related methylation profiles or polymorphisms is useful, for example, in forensic medicine and testing of age of animals in the livestock industry.

The methylation profiling assay of the present invention is particularly useful, therefore, in monitoring aging of cells and may be used in conjunction with pallative care or in clinical management of individuals having the methylation profile of a cell or group of cells at a particular stage of development. This is important not only in human and veterinary health but also in monitoring plants and trees, especially those generated through tissue culture.

Another important application is in the high throughput screening of agents which are capable of demethylating genomes. This may be important, for example, in de-differentiating cells.

In relation to these embodiments, genetic constructs such as used, for example, in gene therapy, can be engineered to attach cell-specific methylation modulation signals to ensure that genes are regulated in the appropriate manner.

In the livestock industry, such as meat production, wool production and breeding programs, amongst many other facets, monitoring methylome profiles in relation to environmental conditions (e.g. drought, weather, change in diet, etc.) or following vaccination programs will be important for ensuring there is no loss of quality or decrease in production levels. This is also important for general toxicity testing of humans and animals. Many disease conditions are mediated by toxins and environmental contaminants and AMP technology provides a means for monitoring for changes in methylome profile.

Accordingly, another aspect of the present invention contemplates a method for detecting a change in a cell's or group of cells' developmental state or a cell's or group of cells' exposure to an internal or external stimulus, said method comprising detecting a change in methylation profile over time by the method comprising obtaining a sample of genomic DNA from said cell or group of cells, digesting a sub-sample of said sample of genomic DNA with a methylation-sensitive enzyme having a recognition nucleotide sequence comprising a nucleotide capable of being methylated at one or more of the above-mentioned sites, subjecting said digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a

cleavable nucleotide sequence of the methylation-sensitive enzyme recognition sequence and subjecting the products of the amplification reaction to separation or other detection means relative to a control but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of the amplification reaction to the separation or other detection means, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by the methylation-sensitive enzyme, wherein the presence of amplification products in enzyme digested and non-digested samples is indicative of a methylation-sensitive enzyme-digestion-resistant marker, and wherein the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-sensitive marker and wherein the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-dependent marker.

More particularly, the present invention provides a method for detecting a change in a cell's or group of cells' developmental state or a cell's or group of cells' exposure to an internal or external stimulus, said method comprising detecting a change in methylation profile over time by the method comprising obtaining a sample of genomic DNA from said cell or group of cells, digesting a sub-sample of said sample of genomic DNA with *Hpa*II which has a recognition nucleotide sequence corresponding to or within said sites, subjecting said digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a *Hpa*II cleavable nucleotide sequence and subjecting the products of the amplification reaction to separation and/or detection means relative to a control but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of the amplification reaction to the separation or other detection means, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by *Hpa*II wherein the presence of amplification products in enzyme digested and non-digested samples is indicative of a *Hpa*II-digestion-resistant marker (H^r), the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a *Hpa*II-digestion-sensitive marker (H^s) and the presence and absence of

amplification products in enzyme digested and undigested samples, respectively, is indicative of a *Hpa*II-digestion-dependent marker (H^d).

In a further embodiment, the present invention provides a method for detecting a change in a cell's or group of cells' physiological status such as from a normal state to a disease state, said method comprising obtaining a sample of genomic DNA from said cell or group of cells, digesting a sub-sample of said sample of genomic DNA with a methylation-sensitive enzyme having a recognition nucleotide sequence comprising a nucleotide capable of being methylated at one or more of the above-mentioned sites, subjecting said digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a cleavable nucleotide sequence of the methylation-sensitive enzyme recognition sequence and subjecting the products of the amplification reaction to separation or other detection means relative to a control but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of the amplification reaction to the separation or other detection means, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by the methylation-sensitive enzyme, wherein the presence of amplification products in enzyme digested and non-digested samples is indicative of a methylation-sensitive enzyme-digestion-resistant marker, and wherein the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-sensitive marker and wherein the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-dependent marker.

In yet another embodiment, the present invention provides a method for detecting a change in a cell's or group of cells' physiological status such as from a normal state to a disease state, said method comprising detecting a change in methylation profile over time by the method comprising obtaining a sample of genomic DNA from said cell or group of cells, digesting a sub-sample of said sample of genomic DNA with *Hpa*II which has a recognition nucleotide sequence corresponding to or within said sites, subjecting said

digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a *Hpa*II cleavable nucleotide sequence and subjecting the products of the amplification reaction to separation or other detection means relative to a control but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of the amplification reaction to the separation or other detection means, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by *Hpa*II wherein the presence of amplification products in enzyme digested and non-digested samples is indicative of a *Hpa*II-digestion-resistant marker (H^r), the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a *Hpa*II-digestion-sensitive marker (H^s) and the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a *Hpa*II-digestion-dependent marker (H^d).

Preferably, the amplification reaction is PCR.

Disease conditions associated with methylation include but are not limited to Rett syndrome, ICF and Fragile X syndrome. Furthermore, the change in methylation profile can correlate with the change in expression patterns of particular genes and, hence, the assay of the present invention is a useful way of identifying new genes (or promoters or other genetic sequences) or for functionalizing a known gene or genetic sequence).

A change in methylome profile over time means that at least two time points are measured and there is a change in methylation profile over at least the two time points.

The present invention further permits the identification of genes and promoters having CpG or CpNpG islands which have *Hpa*II sites or other methylation-sensitive restriction sites. Furthermore, application of the AMP technology permits identification of Class III markers (i.e. H^d sites). The identification of these sites permits identification of potential regulatory regions which can be targeted for agonists or antagonists of gene expression.

The present invention further contemplates mapping methylome profiles with phenotypic and/or genotypic profiles. A database of methylomes can then be produced relative to a phenotypic trait or a particular genotype. Such a database is useful in bioinformatics to establish likely phenotypes of, for example, animal, mammalian, avian, insect, reptillian, amphibian, fungal, yeast, nematode or other cells of a eukaryotic organism. Furthermore, the database may also be used to map junctions between methylated and non-methylated DNA and their association with important biological phenotypes.. The latter frequently show high levels of cell specificity. The present invention further contemplates any method for monitoring these junctions and the association of these junctions with important biological phenotypes.

Methylation profiles are also important in plant tissue culture and/or in recombinant protocols applied to plants.

Thus, in another aspect, the invention contemplates a computer program product for assessing the likely phenotype of a cell based on methylome profile, said product comprising:-

(1) code that receives an input value for one or more of features wherein said features are selected from:-

(a) absence or presence of AMP marker Class I;
(b) absence or presence of AMP marker Class II:
(c) absence or presence of AMP marker Class III;
(d) absence or presence of phenotype mapped to a Class I marker;
(e) absence or presence of phenotype mapped to a Class II marker;
(f) absence or presence of phenotype mapped to a Class III marker; and
(g) absence or presence of a junction between methylated and unmethylated DNA; and
(h) absence or presence of phenotype mapped to a junction between methylated and unmethylated DNA.

(2) a computer readable medium that stores the code.

Yet another aspect of the invention extends to a computer system for assessing the likely phenotype of a cell based on methylation profile wherein said computer system comprises:-

(1) a machine-readable data storage medium comprising a data storage material encoded with machine-readable data, wherein said machine-readable data comprise values for one or more features, wherein said features are selected from:-

(a) absence or presence of AMP marker Class I;
(b) absence or presence of AMP marker Class II:
(c) absence or presence of AMP marker Class III;
(d) absence or presence of phenotype mapped to a Class I marker;
(e) absence or presence of phenotype mapped to a Class II marker;
(f) absence or presence of phenotype mapped to a Class III marker;
(g) absence or presence of a junction between methylated and unmethylated DNA; and
(h) absence or presence of phenotype mapped to a junction between methylated and unmethylated DNA;

(2) a working memory for storing instructions for processing said machine-readable data;

(3) a central-processing unit coupled to said working memory and to said machine-readable storage medium, for processing said machine-readable data to provide comparison of phenotype and AMP marker classes; and

(4) an output hardware coupled to said central processing unit for receiving said data of comparison.

The present invention further extends to searching for homologous AMP markers, once identified. This then provides a basis for identifying similar control mechanisms in other genes, promoters or regulatory regions. For example, tissue-specific AMP markers or AMP markers associated with a disease are cloned, sequenced and mapped on the genome. Coding sequences adjacent or limited to the mapped AMP marker are identified. Genome-wide microarray analysis is conducted with cDNA probes from different tissues or from normal *versus* diseased cells to reveal other coding sequences that are coordinately regulated with the coding sequence linked to the said tissue or disease-specific AMP marker. Computer algorithms are then used to identify conserved nucleotide signatures common to all coordinately expressed genes, which in turn are correlated to the methylation status of an AMP marker in the vicinity of each coding sequence. The discovered nucleotide signature can then be used to discover further coordinately regulated coding sequences around the genome regulated by methylation status of an adjacent AMP marker.

A combined AMP analysis approach, or AMP analysis on individuals or tissues of individuals, is adopted to identify DNA methylation markers linked or associated with important phenoypes such as disease, a diseased tissue (e.g. cancer), aging or exposure to a toxin (e.g. excessive alcohol or smoke). For the combined AMP analysis approach, DNA is bulked for normal *versus* affected individuals or tissues. AMP analysis and comparison between normal and affected DNA bulks identifies candidate diagnostic markers associated with the phenotype. Promising markers identified by this approach are then tested on a large number of normal *versus* affected individuals or tissues. When a marker is identified to be always or almost always or very often associated with the phenotype, the marker will be cloned and sequenced, and a diagnostic test developed. The AMP is also be mapped on the genome to identify coding or non-coding genes linked to the methylation change in affected individuals or tissues. These linked coding or non-coding genes or the encoded proteins are identified as therapeutic targets for the development of drugs or therapy.

In cases where the gene is methylated and silenced in affected individuals or tissues, compounds are screened in high throughput fashion in stable cell lines or individuals to identify drugs that result in demethylation and reactivation of the affected gene. Alternatively, a normal active copy of the affected gene are transfected as a transgene into cells to correct the defect. Such transgenes are introduced with modulating sequences that protect the transgene from methylation and keep it unmethylated and transcriptionally active.

In cases where the gene is unmethylated and transcriptionally active or transcriptionally over-active in affected individuals or tissues, compounds are screened in high throughput fashion in stable cell lines to identify drugs that result in methylation and silencing of the affected gene. Alternatively, a transgene encoding a double stranded RNA homologous to the affected AMP or homologous to sequences linked to the AMP, or double stranded RNA homologous to the AMP or homologous to sequences linked to the AMP, are transfected as a transgene into cells to methylate the gene, silence it and thereby correct the defect. Such double stranded RNA-encoding transgenes are introduced with modulating sequences which protect it from methylation, keep it transcriptionally active and producing double stranded RNA.

The present invention is further described by the following non-limiting Examples.

EXAMPLE 1

AMP protocol

The following protocol may need to be adapted for use in particular plant or animal cells. However, in general, the procedure involves isolating genomic DNA, subjecting the DNA to *Hpa*II digestion and subjecting the digest to an amplification reaction using a primer which spans

all or part of the *Hpa*II recognition sequence, i.e. CCGG, such that if *Hpa*II has digested the nucleotide sequence, primer extension cannot occur.

A diagram showing the AMP protocol is shown in Figure 1.

Three classes of product are expected:-

Class I : *Hpa*II digestion-resistant (H^r) markers, indicative of methylation

Class II : *Hpa*II digestion-sensitive (H^s) markers, indicative of non-methylation .

Class III : *Hpa*II digestion-dependent (H^d) markers. Amplicon is methylated or partly methylated and is linked to an unmethylated *Hpa*II site followed by an inhibitor sequence. Cleavage of the unmethylated site removes inhibition and allows amplification.

The existence of an H^d (Class III) marker is further shown in Figure 2. An H^d marker represents a junction between methylated and unmethylated DNA. Figure 3 provides an example of the different classes of AMP markers.

The following protocol is useful for most cells such as from humans, sugarcane and wheat.

The DNA marker protocols are based on arbitrarily-primed PCR, using high concentrations (5 μM) of small primers, ten nucleotides in length. PCR is performed with a hot start (85°C), an initial denaturation at 94°C for 5 min, followed by 30 cycles of: 94°C for 30 secs, 60 sec at each of 57°C, 56°C, 55°C, 54°C and 53°C and a final extension step at 72°C for 5 min. Highly reproducible RAFs were amplified from either purified DNA or alkali treated leaf tissue (Klimyuk *et al., Plant J. 3:* 493-494, 1993) to demonstrate DNA sequence change between sugarcane lines. The amplified DNA methylation polymorphisms (AMP) protocol is similar to the RAF protocol, but it uses genomic DNA digested with methylation sensitive enzymes before it is used as template in PCR using oligonucleotide primers with recognition sequences for the methylation-sensitive enzyme.

The protocol useful for eukaryotic cells such as from humans, sugarcane or wheat is as follows.

The method is a PCR-based DNA marker protocol for detecting AMP markers in genomic DNA. AMP markers are detected by using oligonucleotide primers carrying recognition sequences for methylation-sensitive restriction endonucleases on genomic template digested with the corresponding enzyme. The procedure listed below is for use of the methylation-sensitive enzyme *Hpa*II and corresponding *Hpa*II oligonucleotide primers, but the protocol can be modified accordingly for all other methylation-sensitive enzymes.

Template preparation:

1. Digest 5.0 μg of high quality genomic DNA with the methylation sensitive enzyme (e.g. *Hpa*II) [overnight digestion is convenient].

Digest Reaction (100 µL)

10 x enzyme buffer	10 µL
*Hpa*II (10 U / µl)	5 µL
DNA	5 µg
sterile dH_2O to final Vol.	100 µL

2. Run 250 ng (5 µL) of digest on 0.7% w/v agarose gel to check digestion.

3. Remove 250 ng (5 µL) and keep for later (to compare with recovery after clean-up).

4. Precipitate digested DNA with 0.1 Vol. 3 M Na Acetate (pH 6), 2 Vol. 100% v/v EtOH, 1 µL (20 µg/µl) glycogen. Mix gently, leave on bench for 10 to 15 mins.

5. Microfuge at 4°C (13000 rpm) 30 mins. Immediately after spinning, remove the supernatant and wash pellet with 70% v/v EtOH. Remove EtOH and dry pellet (2 mins in Speedie Vac or air-dry at room temp in laminar flow).

6. Resuspend pellet in 50 µL sterile dH_2O (ca. 100 ng/µL). To check recovery of the clean-up run 2.5 µL (ca. 250 ng) on gel alongside the 250 ng kept from step 3. Estimate recovery and dilute samples to 50 ng/µL for use in the AMP reaction.

A PCR procedure is then conducted using the following primers:-

***Hpa*II Oligonucleotide Primers (Operon)**		
Oligo	**5'-3'**	**SEQ ID NO:**
B-16	TTTGCCCGGA	1
C-08	TGGACCGGTG	2
D-20	ACCCGGTCAC	3
F-17	AACCCGGGAA	4
F-18	TTCCCGGGTT	5

*Hpa*II Oligonucleotide Primers (Operon)		
Oligo	**5'-3'**	**SEQ ID NO:**
I-08	TTTGCCCGGT	6
J-01	CCCGGCATAA	7
J-14	CACCCGGATG	8
M-17	TCAGTCCGGG	9
N-09	TGCCGGCTTG	10
P-05	CCCCGGTAAC	11
V-15	CAGTGCCGGT	12
V-17	ACCGGCTTGT	13
W-03	GTCCGGAGTG	14
W-15	ACACCGGAAC	15
AB-16	CCCGGATGGT	16
AE-11	AAGACCGGGA	17
AF-16	TCCCGGTGAG	18
AJ-15	GAATCCGGCA	19
AK-18	ACCCGGAAAC	20
AM-09	TGCCGGTTCA	21
AN-14	AGCCGGGTAA	22
AP-20	CCCGGATACA	23
AQ-16	CCCGGAAGAG	24
AR-18	CTACCGGCAC	25
AT-10	ACCTCCGGTC	26
AV-14	CTCCGGATCA	27
AY-03	TTTCCGGGAG	28
BB-09	AGGCCGGTCA	29
BB-18	CAACCGGTCT	30
BF-13	CCGCCGGTAA	31
BG-17	TCCGGGACTC	32

Useful PCR stock solutions include:-

1. 10 x DAF buffer = 100 mM Tris pH 8, 100 mM KCl, 50 mM $MgCl_2$;
2. 1 mM dNTPs;
3. α-labelled ^{33}P-dATP, 10 μCi/μL (as from supplier); and
4. 10 μM 10-mer primer stocks.

The PCR procedure is conducted as follows:-

PCR is conducted in a 10 µL volume, with a single oglionucleotide primer, 10 nucleotides long (see above).

Reaction Mix	1 reaction	10 reactions	Final concentration
Stoffel fragment 10 U/µL	0.15 µL	1.5 µL	1.5 U/10 µL
10 x DAF buffer	1.0 µL	10.0 µL	
dNTPs (1 mM)	0.2 µL	2.0 µL	20 µM
Sterile dH_2O	2.4 µL	24.0 µL	
α-labeled ^{33}P-dATP (10 µCi/µL)	0.25 µL	2.5 µL	2.5 µCi/10 µL
Total	**4.0 µL**	**40.0 µL**	

Aliquot into each PCR tube 4 µL of reaction mix, then add 5 µL primer (final concentration 5 µM) and 1 µl 50 ng/µL digested or genomic DNA.

Prepare reactions on ice and load tubes into the PCR machine with the block at 85°C (i.e. hot start).

The following conditions are employed:-

1.	94°C	5 min
2.	94°C	30 sec
3.	57°C	1 min
4.	56°C	1 min
5.	55°C	1 min
6.	54°C	1 min
7.	53°C	1 min
8.	go to step (2)	29 more times
9.	72°C	5 min

10. end.

The protocol was slightly modified for the large wheat genome, in that the annealing/extension temperatures were 2°C higher, i.e.

30 cycles of	94°C for 5 min
	94°C for 30 sec
	59°C for 1 min
	58°C for 1 min
	57°C for 1 min
	56°C for 1 min
	55°C for 1 min.

Furthermore, the following changes were also made:-

1. Add 10 μL of 2x loading buffer (0.0% w/v Bromophenol Blue, 0.05% Xylene Cyanol, 98% v/v formamide, 10 mM EDTA, pH 8) to each PCR sample, mix gently. Before gel electrophoresis, denature PCR reactions at 90°C for 3 mins, cool on ice.

2. Separate the products on sequencing gels (4% w/v polyacrylamide gel, 7.5 M urea in 1x TBE, run at100 W, 2 hr 15 mins). Dry gel and expose to film for 6 hours or overnight.

The extent of DNA methylation change was determined in 24 sugarcane lines cloned by tissue culture. All 24 lines investigated had widespread DNA methylation changes and unique AMP profiles. An average of 0.8% DNA methylation polymorphism over more than 1,000 markers was observed between the parent variety and somaclones and both increases and decreases in DNA methylation were observed. DNA methylation polymorphism was also demonstrated in conventionally propagated plants, but the level was about two-fold greater ($p < 0.05$) in tissue culture propagated plants.

Clearly, DNA methylation change predisposed sugarcane clones to genome instability as about 20% of non-transgenic somaclones lost DNA as measured by randomly amplified DNA fingerprinting protocols (RAFs) and this finding was confirmed with simple sequence repeat (SSR) markers. Up to 2% of RAF markers were lost in a selection of transgenic lines of sugarcane. Genetic mapping of DNA markers susceptible to loss showed that they were often deleted as a linkage group. The inventors detected an example of chromosome breakage as distinct from chromosome loss in one transgenic line. Interestingly, there was a significant correlation between loss of markers and reduced yield in preliminary field trials on transgenic lines ($p = 0.017$).

This example showed that a minimal passage through tissue culture altered the methylation status of about 1% of the sugarcane genome, as randomly detected by the AMP technology. Transposon activity can lead to chromosome loss and/or breakage (e.g., see English *et al., Plant Cell 5:* 501-514, 1993). Thus, the increased DNA methylation change in somaclones as shown here may have direct effects on gene expression as well as predisposing the sugarcane genome to instability. The AMP technologies in accordance with the present invention may usefully be applied as diagnostic tests to identify tissue culture conditions that minimize DNA methylation change and loss of DNA during the tissue culture phase of genetic engineering of plants and animals.

EXAMPLE 2

Detection of DNA polymorphisms in human body fluid

Using the AMP technology, abundant and distinct tissue-specific DNA methylation polymorphisms were readily detected between human blood and sperm (Figure 4). Most of the tissue specific DNA methylation polymorphisms were Class III digestion dependent markers and surprisingly, 30% of these were polymorphic in methylation status between these tissues. Between five blood samples and three sperm samples (from eight separate individuals), 80% of variation in AMP profiles generated from *Hpa*II-digested template

could be explained by the tissue type, and the remaining 20% of variation could not be distinguished from nucleotide (rather than DNA methylation) polymorphism.

In another example, blood from three sets of young, healthy, identical twins were also assessed for DNA methylation polymorphisms over 8,000 CpG sites in the genome, representing an average of about 340 CpG sites per human chromosome. There was 100% concordance of AMP profiles of identical twins thereby indicating stringent genetic programming of DNA methylation patterns in human cells to produce cell-specific "methylomes". Interestingly, an AMP marker was detected that showed an age-dependent methylation change. This has implications in targets for therapy of age-related conditions or illnesses and in forensic medicine.

EXAMPLE 3

Identification of Class III markers in human genome

Figure 5 is a diagrammatic representation showing map location of four human digestion-dependent Class III AMP markers and the flanking *Hpa*II sites. Each of the digestion-dependent markers that was cloned and sequenced was 100% homologous to sequenced regions of the human genome. The positions of the Class III amplicons (HdM 1-4) [SEQ ID NOS:33-36] are indicated by thick shaded boxes on the chromosomal segment. The locations of *Hpa*II sites are represented by vertical lines and the distance to, and orientation of nearby genes are indicated. Southern analysis has shown that Class III markers represent the junction between methylated and unmethylated DNA in the genome. Tissue-specific Class III AMP markers often map to the CpG islands and often represent tissue-specific modification of methylation in or around CpG islands.

EXAMPLE 4

Identification of promoter elements that contain CpG islands in human genomic DNA

Human promoter elements that contains CpG islands are identified by searching the human genome database and selecting promoter element sequences that comprise regions of DNA which are approximately 180 bp in length and that have a sequence complexity comprising a G+C content above 50%. An example of such a promoter element comprising a CpG island is rpA2 promoter element (shown below):-

```
caggcgcccg ccaccacacc cagctaattt tttatatttt tagtagagat ggggtttcac   60
cgtgttagcc agcatggtct tgatctcctg atctcgtgat cagggaaatc tgtttaatta  120
aaacacgctt tgcgatttct gcccagtgct ctgaatgtca aagtcaagag taggaggaag  180
aagaaagaag aaagcagggt gcagtggctg acgcctgtaa tcccaacact ttgggaggcc  240
aaggtgggtg gatcacttga ggtcaggagt tcgagaccag cctgaccagc agggcgaaac  300
ccccgtctct actaaaaata caaaaattag ccagggcgtg gtggcacgtg ctgtaatccc  360
agctacgccg gaggctaagg caggaaaatc gcttgaaccc gggaggcgga ggttgcagtg  420
aggtgagatc gcaccattgc actccagcct gggcaacaag agcaaaactc tgtctccgaa  480
aaaaaaaaaa aaaggaaaaa gaagagaaga aaaaaacaca cagtcctgga gcccaaaggc  540
catgtctatc ttacttaacc attgaatccc agtggccagc acagggccag gaacaaagaa  600
ggcataaaca attctttttt ttttttttt ttttttttga gtcgcagtct ctctctgcca  660
cctaggctgg agtgcagtgg cagcatcaca gctcactgca gcctggacac cccagggcca  720
ggtgatcctc ccaccttagc caccccagta gctgggaccc caggcacttg ccaccacacc  780
agactaattt ttaaaaatat tttttgagag agggtctcac tatgttgtcc aggctggtct  840
caaacttcca gcctcaagcg gtcctcctgc ctcagacccc atttgctggg tttacaggca  900
tgagccacag cacctgctaa ttttttcttaa atacataaat gaacataaaa ttctaacaat  960
gcatgagtat tttgaggaag gaactgacaa aatgttccac tccctatggg aggcacgtta 1020
tatgaagaat tatgaaaaat ggtcgaaatg actggagagg ccaagcctgg atgagactgg 1080
gatggggaca ggtgcgggac gaggggcacc accctcacat ctttcacaag tctgtcatag 1140
gcaagagggc gtaggtttct cacagcccca ctggggagaa tcggcaccat tggtggcatt 1200
acacgaagag aatgtgacct cctatgtaaa agaacaagca actccacgcg gtgctgtgag 1260
gctagtgctg cgagtccctg aggtgcgcaa ttcccgcacg accgtgggtg ggaaacaccg 1320
aagccaaaac tccgctacag ccctttagat gaaggcgtcg tctgattggt gatagtttgg 1380
cgcgaacctg agcacgccga acaaaggaag tgacggcaga agtcgcgcac ttgacgaggg 1440
tgggatcaca cggcgctgcg tcgcggtagt attgttctga ttggttgatt tcttgcgata 1500
ccgctctgcc agccccttgc ttccgctagt gcggagggtt ttgcccttcg taaagatggc 1560
```

cgcggaggct tttggagcca actgggagcg cagtacgcgt tttctggagc atgggcagag 1620
gagacaggaa caagcgtagc atccgtgagc accgattggc tgaagcgagc accccgggag 1680
ctgactggct ccgccattcg cgggaaggcg tttgtggtgc cagagaaaag tagccagagc 1740
ggcgcagtgg 1750

[SEQ ID NO:37]

The CpG islands present in the promoter element are highlighted. The rpA2 promoter element contains two CpG islands, which are highlighted. CpG islands in this promoter element indicate that the expression of the gene is regulated by the methylation status of the promoter element.

EXAMPLE 5

Detection of promoter elements that contain CpG islands in plant genomic DNA

A plant promoter element that contains CpG islands is identified by searching plant genome databases and selecting promoter element sequences that comprise regions of DNA which are at least 180 bp in length and that have a sequence complexity comprising a G+C content above 50%. An example of such a promoter element comprising a CpG island is the rbp21 promoter element derived from the rbp21 gene from *Oryza sativa* (shown below):

ctgcagagag gatgaccctt gtcaccaccg tcatgtacga ggctgcttca ccactgcctc 60
actgccacca gcgtctcccg ccgcgtgcaa tacaagaaga aacatcgaac ggtcatataa 120
ggtaagaccc actaccgatt taacctatca ttcccacaat ctaatccact tatttctctt 180
cccatgatct tatcctctca tttctcctca ctacttttgc atttgtagga aacacaatga 240
caccgtcgaa gaaagctggt ggagcaccgt agccagcaat caccaaaaca cagaggggag 300
gaggtcggca gcggccatgc ggacggcgat gagacaacgc gacgcaaaga gggaggagga 360
cgttggcgat catgctggtg ttggcggagg aggtcactgg ccatgcgaat gacagcgggg 420
cagcgcaaca caaaaagggg ggaggatgcc ggcgaccacg ctagtaccat gaagcaagat 480
gatgtgaaag ggaggaccgg acgagggttg gacctctgcc gccgacgtga agagcgtgat 540
gtgtagaagg agatgttaga ccagatgccg acgcaactta gccctgcaag tcacccgact 600
gcatatcgct gcttgccctc gtcctcatgt acacaatcag cttgcttatc tctccatact 660
tgtcgtttgt ttcccgtggc cgaaatagaa gaagacagag gtgggttttg ttggagagtt 720
ttagtggtat tgtaggccta tttgtaattt tgttgtactt tattgtatta atcaataaag 780
gtgtttcatt ctattttgac tcaatgttga atccattgat ctcttggtgt tgcactcagt 840

atgttagaat attcattccg ttgaaacaat cttggttaag ggttggaaca tttttatctg 900
ttcggtgaaa catccgtaat attttcgttg aaacaatttt tatccgacag caccgtccaa 960
caatttacac caatttggac gtgtgataca tagcagtccc caagtgaaac tgaccaccag 1020
ttgaaaggta tacaaagtga acttattcat ctaaaagacc gcagagatgg gccgtggccg 1080
tggctgcgaa acgacagcgt tcaggcccat gagccattta ttttttaaaa aaatatttca 1140
acaaaaaaga gaacggataa aatccatcga aaaaaaaaaa ctttcctacg catcctctcc 1200
tatctccatc cacggcgagc actcatccaa accgtccatc cacgcgcaca gtacacacac 1260
atagttatcg tctctccccc cgatgagtca ccacccgtgt cttcgagaaa cgcctcgccc 1320
gacaccgtac gtgcgccacc gccgcgcctg ccgcctggac acgtccggct cctctcccgc 1380
cgcgctggcc accgtccacc ggctcccgca cacgtctccc tgtctccctc cacccatgcc 1440
gtggcaatcg agctcatctc ctcgcctcct ccggcttata aatggcggcc accaccttca 1500
cctgcttgca [SEQ ID NO:38]

The CpG islands present in the promoter element are highlighted. The rbp21 promoter element contains two CpG islands. The presence of a CpG island indicates that the activity of the promoter element is regulated by methylation.

EXAMPLE 6

Use of AMP to assess stable mammalian cell lines

Mammalian cell lines are used for many applications in medical research. It is important in many applications that cell lines used are stable in their phenotypes. AMP is used to identify stable *versus* unstable mammalian cell lines. AMP has been used to demonstrate that subcultures of the human colon carcinoma cell line HCT116 are very stable and identical in terms of DNA methylation and DNA sequence (Figure 8). Similarly, AMP can be used to monitor DNA methylation and genome stability in cloned animals.

EXAMPLE 7

Use of AMP to assess stable cell lines

Cell lines that are stable in terms of DNA methylation and DNA sequence can be used with the AMP protocol to identify and discover genes involved in control of DNA methylation and genome stability, or to identify processes that destabilize DNA

methylation and genome integrity. AMP has been used to demonstrate that knocking out the tumor suppressor gene p53 in the human colon carcinoma cell line HCT116 results in the induction of about 3% DNA methylation and 0.3% DNA sequence change to the genome as measured by AMP (Figure 9). These genomic changes could be due to loss of p53 function (indicating a key role for p53 in control of DNA methylation and maintenance of genome stability) or to the physical process of transfecting the cell line. Thus AMP can be used to identify and discover genes involved in control of DNA methylation and genome stability, and to identify processes or toxins that destabilize DNA methylation and genome integrity in mammals.

EXAMPLE 8

Detection of DNA methylation polymorphisms in genomic DNA

Southern analysis confirms that AMP detects DNA methylation polymorphisms in genomic DNA and the junctions of methylated and unmethylated DNA in the genome (Figure 10).

EXAMPLE 9

Combined AMP analysis

DNA from normal or abnormal tissue (e.g. cancer) is separately bulked. AMP analysis is conducted to compare DNA from the normal and abnormal tissue. Candidate fragments present or absent in abnormal tissue relative to normal tissue are then cloned and analyzed for subsequent development as a potential diagnostic marker for disease and/or as a potential therapeutic product itself or as a target for antagonists and agonists. In addition, methylation- or demethylation-facilitating sequences may be used to methylate or demethylate target DNA.

EXAMPLE 10

Use of substractive hybridization to enrich for HpaII dependent markers

With comparing genomes subjected to AMP analysis, given the large number of bands which are present, it is possible that *Hpa*II digestion-dependent markers are masked by other bands. Subtractive amplification is used to increase the number of *Hpa*II-digestion dependent PCR detected in each PCR.

The PCR-Select cDNA subtraction kit (Clontech, Palo Alto, USA) is conveniently used for the subtractive hybridization.

Those skilled in the art will appreciate that the invention described herein is susceptible to variations and modifications other than those specifically described. It is to be understood that the invention includes all such variations and modifications. The invention also includes all of the steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations of any two or more of said steps or features.

BIBLIOGRAPHY

Baylin, S. B. and Herman, J. G. (2000) *Trends Genet. 16:* 168-174.

Bentley, S. and Bassam, B. J. (1996). A robust DNA amplification fingerprinting system applied to analysis of genetic variation within Fusarium oxysporum fsp cubense. *J. Phytopath. 14:* 207-214.

Bonner and Laskey (1974) *Eur. J. Biochem. 46:* 83.

English, J., Harrison, K. and Jones, J. D. (1993). A genetic analysis of DNA sequence requirements for Dissociation state I activity in tobacco. *Plant Cell 5:* 501-14.

Hendrich, B. (2000) *Current Biology 10:* 60-63.

Herman, J. G. (1999) *Semin. Cancer Biol. 9:* 359-67.

Klimyuk, V. I., Carroll, B. J., Thomas, C. M and Jones, J. D. (1993). Alkali treatment for rapid preparation of plant material for reliable PCR analysis. *Plant J. 3:* 493-494.

Laird, P. W. and Jaenisch, R. (1996) *Annu. Rev. Genet. 30:* 441-464.

Lengauer, C. *et al.* (1997) *Proc. Natl. Acad. Sci. USA. 94:* 2545-2550.

Marmur and Doty (1962) *J. Mol. Biol. 5:* 109.

O'Neill, R.J.W. *et al.* (1998) *Nature 393:* 68-72.

Walsh, C. P. and Bestor, T. H. (1999) *Genes Dev. 13*: 26-34.

Yoder, J.A. *et al.* (1997) *Trends Genet. 13:* 335-340.

CLAIMS

1. A method for determining the methylation profile of one or more nucleotides at one or more sites within the genome of a eukaryotic cell or group of cells, said method comprising obtaining a sample of genomic DNA from said cell or group of cells, digesting a sub-sample of said sample of genomic DNA with a methylation-sensitive enzyme having a recognition nucleotide sequence comprising a nucleotide sequence capable of methylation at one or more of the above-mentioned sites, subjecting said digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a cleavable nucleotide sequence of the methylation-sensitive enzyme recognition sequence and subjecting the products of the amplification to separation or other detection means relative to a control, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by the methylation-sensitive enzyme but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of this amplification reaction to said separation or other detection means, wherein the presence of the same amplification products in enzyme digested and non-digested samples is indicative of a methylation-sensitive enzyme-digestion-resistant marker, the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-sensitive markers, and the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-dependent marker.

2. The method of Claim 1 wherein the junctions of methylated and unmethylated DNA are detected.

3. The method of Claim 1 wherein the potentially-methylated sites are CpG or CpNpG nucleotides within the recognition sequence of the methylation-sensitive enzyme.

4. The method of Claim 1 or 2 or 3 wherein the methylation-sensitive enzyme is selected from the list comprising *Aat*II, *Aci*I, *Acl*I, *Age*I, *Asc*I, *Ava*I, *Bam*HI, *Bsa*A1,

*Bsa*H1, *Bsi*E, *Bsi*W, *Bsr*F, *Bss*HII, *Bst*BI, *Bst*UI, *Cla*1, *Eag*I, *Hae*II, *Hga*I, *Hha*I, *Hin*PI, *Hpa*II, *Mlo*I, *Msp*I, *Nae*I, *Nar*I, *Not*I, *Nru*I and *Pml*I.

5. The method of Claim 4 wherein the methylation-sensitive enzyme is *Hpa*II or its functional equivalent.

6. The method of Claim 5 wherein the methylation-sensitive enzyme is *Hpa*II.

7. The method of Claim 1 wherein the amplification reaction is selected from polymerase chain reaction (PCR), ligase chain reaction (LCR), gap filling LCR (GLCR), Qβ replicase; stand displacment amplification (SDA); self-sustained sequence replication (3SR) nucleic acid sequence-based amplification (NASBA).

8. The method of Claim 7 wherein the amplification reaction is PCR.

9. The method of Claim 8 wherein the amplification reaction is real-time PCR.

10. The method of Claim 1 wherein the separation means is gel electrophoresis.

11. The method of Claim 1 wherein the eukaryotic cell is a human or mammalian embryonic stem (ES) cell, embryonic germ (EG) cell, post-natal stem cell, committed cell or mature cell or a cell having a state of development in between any of the above cells.

12. The method of Claim 1 wherein the eukaryotic cell is a cancer cell or cancer cell line.

13. The method of Claim 1 wherein the cell is a plant cell.

14. The method of Claim 1 wherein the cell is a non-mammalian animal cell.

15. The method of Claim 1 wherein the cell is a mammalian cell.

16. A method for determining the methylation profile of one or more CpG or CpNpG nucleotides at one or more sites within the genome of a eukaryotic cell or group of cells, said method comprising obtaining a sample of genomic DNA from said cell, digesting a sub-sample of said sample of genomic DNA with *Hpa*II which has a recognition nucleotide sequence corresponding to or within said sites, subjecting said digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a *Hpa*II cleavable nucleotide sequence and subjecting the products of the amplification reaction to separation or other detection means relative to a control but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of the amplification reaction to the separation or other detection means, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by *Hpa*II, wherein the presence of amplification products in enzyme digested and non-digested samples is indicative of a *Hpa*II-digestion-resistant marker (H^r), the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a *Hpa*II-digestion-sensitive marker (H^s) and the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a *Hpa*II-digestion-dependent marker (H^d).

17. The method of Claim 16 wherein the junctions of methylated and unmethylated DNA are detected.

18. The method of Claim 16 or 17 wherein the amplification reaction is PCR.

19. The method of Claim 18 wherein the amplification reaction is real-time PCR.

20. The method of Claim 16 wherein the eukaryotic cell is a human or mammalian embryonic stem (ES) cell, embryonic germ (EG) cell, post-natal stem cell, committed cell or mature cell or a cell having a state of development in between any of the above cells.

21. The method of Claim 16 wherein the eukaryotic cell is a cancer cell or a cancer cell line.

22. The method of Claim 16 wherein the cell is a plant cell.

23. The method of Claim 16 wherein the cell is a non-mammalian animal cell.

24. The method of Claim 14 wherein the cell is a mammalian cell.

25. A methylation profile of the genome of a eukaryotic cell or group of cells, the methylation profile comprising the presence or absence of methylation at particular sites and/or junctions between methylated and unmethylated regions, said method comprising obtaining a sample of genomic DNA from said cell or group of cells, digesting a sub-sample of said sample of genomic DNA with a methylation-sensitive enzyme having a recognition nucleotide sequence comprising a nucleotide capable of methylation at one or more of the above-mentioned sites, subjecting said digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a cleavable nucleotide sequence of the methylation-sensitive enzyme recognition sequence and subjecting the products of the amplification reaction to separation or other detection means relative to a control but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of the amplification reaction to the separation or other detection means, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by the methylation-sensitive enzyme wherein the presence of amplification products in enzyme digested and non-digested samples is indicative of a methylation-sensitive enzyme-digestion-resistant marker, the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-sensitive marker and the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-dependent marker.

26. The method of Claim 25 wherein the junctions of methylated and unmethylated DNA are detected.

27. The method of Claim 25 or 26 wherein the potentially-methylated sites are CpG or CpNpG nucleotides within the recognition sequence of the methylation-sensitive enzyme.

28. The method of Claim 25 or 26 or 27 wherein the methylation-sensitive enzyme is selected from the list comprising *Aat*II, *Aci*I, *Acl*I, *Age*I, *Asc*I, *Ava*I, *Bam*HI, *Bsa*A1, *Bsa*H1, *Bsi*E, *Bsi*W, *Bsr*F, *Bss*HII, *Bst*BI, *Bst*UI, *Cla*1, *Eag*I, *Hae*II, *Hga*I, *Hha*I, *Hin*PI, *Hpa*II, *Mlo*I, *Msp*I, *Nae*I, *Nar*I, *Not*I, *Nru*I and *Pml*I.

29. The method of Claim 28 wherein the methylation-sensitive enzyme is *Hpa*II or its functional equivalent.

30. The method of Claim 29 wherein the methylation-sensitive enzyme is *Hpa*II.

31. The method of Claim 25 wherein the amplification reaction is selected from polymerase chain reaction (PCR), ligase chain reaction (LCR), gap filling LCR (GLCR), Qβ replicase; stand displacment amplification (SDA); self-sustained sequence replication (3SR) nucleic acid sequence-based amplification (NASBA).

32. The method of Claim 31 wherein the amplification reaction is PCR.

33. The method of Claim 32 wherein the amplification reaction is real-time PCR.

34. The method of Claim 25 wherein the separation means is gel electrophoresis.

35. The method of Claim 25 wherein the eukaryotic cell is a human or mammalian embryonic stem (ES) cell, embryonic germ (EG) cell, post-natal stem cell, committed cell or mature cell or a cell having a state of development in between any of the above cells.

36. The method of Claim 25 wherein the eukaryotic cell is a cancer cell or cancer cell line.

37. The method of Claim 25 wherein the cell is a plant cell.

38. The method of Claim 25 wherein the cell is a non-mammalian animal cell.

39. The method of Claim 21 wherein the cell is a mammalian cell.

40. An assay device in the form of a kit useful for determining the methylation profile of one or more nucleotides at one or more sites within the genome of a eukaryotic cell, said kit comprising in compartmental form multiple compartments each adapted to comprise one or more of buffers, diluents and enyzmes as single or multiple components which are optionally required to be admixed prior to use, said kit further comprising instructions for use, wherein the method is conducted by obtaining a sample of genomic DNA from said cell, digesting a sub-sample of said sample of genomic DNA with a methylation-sensitive enzyme having a recognition nucleotide sequence comprising a nucleotide capable of being methylated at one or more of the above-mentioned sites, subjecting said digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a cleavable nucleotide sequence of the methylation-sensitive enzyme recognition sequence and subjecting the products of the amplification reaction to separation or detection means relative to a control but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of the amplification reaction to the separation or other detection means, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by the methylation-sensitive enzyme, wherein the presence of amplification products in enzyme digested and non-digested samples is indicative of a methylation-sensitive enzyme-digestion-resistant marker, and wherein the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-sensitive

marker and wherein the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-dependent marker.

41. The kit of Claim 40 wherein the junctions of methylation and unmethylated DNA are detected.

42. The kit of Claim 40 wherein wherein the potentially-methylated sites are CpG or CpNpG nucleotides within the recognition sequence of the methylation-sensitive enzyme.

43. The kit of Claim 40 or 41 or 42 wherein the methylation-sensitive enzyme is selected from the list comprising *Aat*II, *Aci*I, *Acl*I, *Age*I, *Asc*I, *Ava*I, *Bam*HI, *Bsa*A1, *Bsa*H1, *Bsi*E, *Bsi*W, *Bsr*F, *Bss*HII, *Bst*BI, *Bst*UI, *Cla*1, *Eag*I, *Hae*II, *Hga*I, *Hha*I, *Hin*PI, *Hpa*II, *Mlo*I, *Msp*I, *Nae*I, *Nar*I, *Not*I, *Nru*I and *Pml*I.

44. The kit of Claim 43 wherein the methylation-sensitive enzyme is *Hpa*II or its functional equivalent.

45. The kit of Claim 44 wherein the methylation-sensitive enzyme is *Hpa*II.

46. The kit of Claim 40 wherein the amplification reaction is selected from polymerase chain reaction (PCR), ligase chain reaction (LCR), gap filling LCR (GLCR), Qβ replicase; stand displacment amplification (SDA); self-sustained sequence replication (3SR) nucleic acid sequence-based amplification (NASBA).

47. The kit of Claim 46 wherein the amplification reaction is PCR.

48. The kit of Claim 47 wherein the amplification reaction is real-time PCR.

49. The kit of Claim 40 wherein the separation means is gel electrophoresis.

50. The kit of Claim 40 wherein the eukaryotic cell is a human or mammalian embryonic stem (ES) cell, embryonic germ (EG) cell, post-natal stem cell, committed cell or mature cell or a cell having a state of development in between any of the above cells.

51. The kit of Claim 40 wherein the eukaryotic cell is a cancer cell or cancer cell line.

52. The kit of Claim 40 wherein the cell is a plant cell.

53. The kit of Claim 40 wherein the cell is a non-mammalian animal cell.

54. The kit of Claim 40 which detects the junctions of methylated and unmethylated DNA in cells.

55. A method for detecting a change in a cell's or group of cells' developmental state or a cell's or group of cells' exposure to an internal or external stimulus, said method comprising detecting a change in methylation profile over time by the method comprising obtaining a sample of genomic DNA from said cell, digesting a sub-sample of said sample of genomic DNA with a methylation-sensitive enzyme having a recognition nucleotide sequence comprising a nucleotide capable of being methylated at one or more of the above-mentioned sites, subjecting said digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a cleavable nucleotide sequence of the methylation-sensitive enzyme recognition sequence and subjecting the products of the amplification reaction to separation or other detection means relative to a control but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of the amplification reaction to the separation or other detection means, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by the methylation-sensitive enzyme, wherein the presence of amplification products in enzyme digested and non-digested samples is indicative of a methylation-sensitive enzyme-

digestion-resistant marker, and wherein the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-sensitive marker and wherein the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-dependent marker.

56. The method of Claim 55 wherein the junctions of methylated and unmethylated DNA are detected.

57. The method of Claim 55 wherein the potentially-methylated sites are CpG or CpNpG nucleotides within the recognition sequence of the methylation-sensitive enzyme.

58. The method of Claim 55 or 56 or 57 wherein the methylation-sensitive enzyme is selected from the list comprising *Aat*II, *Aci*I, *Acl*I, *Age*I, *Asc*I, *Ava*I, *Bam*HI, *Bsa*A1, *Bsa*H1, *Bsi*E, *Bsi*W, *Bsr*F, *Bss*HII, *Bst*BI, *Bst*UI, *Cla*1, *Eag*I, *Hae*II, *Hga*I, *Hha*I, *Hin*PI, *Hpa*II, *Mlo*I, *Msp*I, *Nae*I, *Nar*I, *Not*I, *Nru*I and *Pml*I.

59. The method of Claim 58 wherein the methylation-sensitive enzyme is *Hpa*II or its functional equivalent.

60. The method of Claim 59 wherein the methylation-sensitive enzyme is *Hpa*II.

61. The method of Claim 55 wherein the amplification reaction is selected from polymerase chain reaction (PCR), ligase chain reaction (LCR), gap filling LCR (GLCR), Qβ replicase; stand displacment amplification (SDA); self-sustained sequence replication (3SR) nucleic acid sequence-based amplification (NASBA).

62. The method of Claim 61 wherein the amplification reaction is PCR.

63. The method of Claim 62 wherein the amplification reaction is real-time PCR.

64. The method of Claim 61 wherein the separation means is gel electrophoresis.

65. The method of Claim 61 wherein the eukaryotic cell is a human or mammalian embryonic stem (ES) cell, embryonic germ (EG) cell, post-natal stem cell, committed cell or mature cell or a cell having a state of development in between any of the above cells.

66. The method of Claim 61 wherein the eukaryotic cell is a cancer cell or cancer cell line.

67. The method of Claim 16 wherein the cell is a plant cell.

68. The method of Claim 61 wherein the cell is a non-mammalian animal cell.

69. A computer program product for assessing the likely phenotype of a cell based on methylome profile, said product comprising:-

(1) code that receives an input value for one or more of features wherein said features are selected from:-

(a) absence or presence of AMP marker Class I;
(b) absence or presence of AMP marker Class II:
(c) absence or presence of AMP marker Class III;
(d) absence or presence of phenotype mapped to a Class I marker;
(e) absence or presence of phenotype mapped to a Class II marker;
(f) absence or presence of phenotype mapped to a Class III marker;
(g) absence or presence of a junction between methylated and unmethylated DNA; and
(h) absence or presence of phenotype mapped to a junction between methylated and unmethylated DNA;

(2) a computer readable medium that stores the code.

70. A computer system for assessing the likely phenotype of a cell based on methylation profile wherein said computer system comprises:-

(1) a machine-readable data storage medium comprising a data storage material encoded with machine-readable data, wherein said machine-readable data comprise values for one or more features, wherein said features are selected from:-

(a) absence or presence of AMP marker Class I;
(b) absence or presence of AMP marker Class II:
(c) absence or presence of AMP marker Class III;
(d) absence or presence of phenotype mapped to a Class I marker;
(e) absence or presence of phenotype mapped to a Class II marker;
(f) absence or presence of phenotype mapped to a Class III marker;
(g) absence or presence of a junction between methylated and unmethylated DNA; and
(h) absence or presence of phenotype mapped to a junction between methylated and unmethylated DNA;

(2) a working memory for storing instructions for processing said machine-readable data;

(3) a central-processing unit coupled to said working memory and to said machine-readable storage medium, for processing said machine-readable data to provide comparison of phenotype and AMP marker classes; and

(4) an output hardware coupled to said central processing unit for receiving said data of comparison.

71. An isolated stem cell or stem cell line or culture of stem cells characterized as having a methylation profile is determined by the method of Claim 1 or 16.

72. A method for identifying a genetic sequence associated with a disease phenotype or unwanted trait, said method comprising determining the methylation profie within the genomes of normal compared to diseased tissue by obtaining a sample of genomic DNA from said cell or group of cells, digesting a sub-sample of said sample of genomic DNA with a methylation-sensitive enzyme having a recognition nucleotide sequence comprising a nucleotide sequence capable of methylation at one or more of the above-mentioned sites, subjecting said digested DNA to an amplification reaction using primers comprising a nucleotide sequence capable of annealing to a non-cleaved form of a cleavable nucleotide sequence of the methylation-sensitive enzyme recognition sequence and subjecting the products of the amplification to separation or other detection means relative to a control, said control comprising another sub-sample of said sample of genomic DNA not subjected to digestion by the methylation-sensitive enzyme but subjected to an amplification reaction using the same primers as for the digested DNA sample and then subjecting the products of this amplification reaction to said separation or other detection means, wherein the presence of the same amplification products in enzyme digested and non-digested samples is indicative of a methylation-sensitive enzyme-digestion-resistant marker, the absence and presence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-sensitive markers, and the presence and absence of amplification products in enzyme digested and undigested samples, respectively, is indicative of a methylation-sensitive enzyme-digestion-dependent marker wherein the presence of particular markers is indicative of normal or diseased tissue.

73. The method of Claim 72 wherein the potentially-methylated sites are CpG or CpNpG nucleotides within the recognition sequence of the methylation-sensitive enzyme.

74. The method of Claim 72 or 73 wherein the methylation-sensitive enzyme is selected from the list comprising *Aat*II, *Aci*I, *Acl*I, *Age*I, *Asc*I, *Ava*I, *Bam*HI, *Bsa*A1,

*Bsa*H1, *Bsi*E, *Bsi*W, *Bsr*F, *Bss*HII, *Bst*BI, *Bst*UI, *Cla*1, *Eag*I, *Hae*II, *Hga*I, *Hha*I, *Hin*PI, *Hpa*II, *Mlo*I, *Msp*I, *Nae*I, *Nar*I, *Not*I, *Nru*I and *Pml*I.

75. The method of Claim 74 wherein the methylation-sensitive enzyme is *Hpa*II or its functional equivalent.

76. The method of Claim 75 wherein the methylation-sensitive enzyme is *Hpa*II.

77. The method of Claim 72 wherein the amplification reaction is selected from polymerase chain reaction (PCR), ligase chain reaction (LCR), gap filling LCR (GLCR), Qβ replicase; stand displacment amplification (SDA); self-sustained sequence replication (3SR) nucleic acid sequence-based amplification (NASBA).

78. The method of Claim 77 wherein the amplification reaction is PCR.

79. The method of Claim 78 wherein the amplification reaction is real-time PCR.

80. A method of treating a subject having a disease state associated with silencing of a genetic sequence as determined by the method of Claim 72, said method comprising introducing into said subject said genetic sequence flanked bymethylation-preventing sequences to thereby replace the silenced genetic sequences or introducing a protein or non-proteinaceous product which replaces the protein encoded by the silenced genetic sequence.

81. The method of Claim 80 wherein the potentially-methylated sites are CpG or CpNpG nucleotides within the recognition sequence of the methylation-sensitive enzyme.

82. The method of Claim 80 or 81 wherein the methylation-sensitive enzyme is selected from the list comprising *Aat*II, *Aci*I, *Acl*I, *Age*I, *Asc*I, *Ava*I, *Bam*HI, *Bsa*A1, *Bsa*H1, *Bsi*E, *Bsi*W, *Bsr*F, *Bss*HII, *Bst*BI, *Bst*UI, *Cla*1, *Eag*I, *Hae*II, *Hga*I, *Hha*I, *Hin*PI, *Hpa*II, *Mlo*I, *Msp*I, *Nae*I, *Nar*I, *Not*I, *Nru*I and *Pml*I.

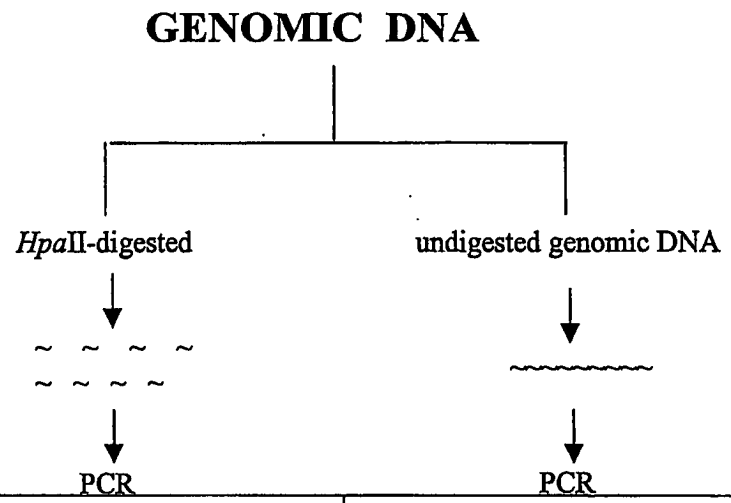
83. The method of Claim 82 wherein the methylation-sensitive enzyme is *Hpa*II or its functional equivalent.

84. The method of Claim 83 wherein the methylation-sensitive enzyme is *Hpa*II.

85. The method of Claim 80 wherein the amplification reaction is selected from polymerase chain reaction (PCR), ligase chain reaction (LCR), gap filling LCR (GLCR), Qβ replicase; stand displacment amplification (SDA); self-sustained sequence replication (3SR) nucleic acid sequence-based amplification (NASBA).

86. The method of Claim 85 wherein the amplification reaction is PCR.

87. The method of Claim 86 wherein the amplification reaction is real-time PCR.



AMP MARKER		
*Hpa*II Sensitive AMP marker (H^s)	(-)	(+)
*Hpa*II Resistant AMP marker (H^r)	(+)	(+)
*Hpa*II Dependant AMP marker (H^d)	(+)	(-)

The symbol in parentheses represents the presence [(+)] or absence [(-)] of an amplicon indicated as a band on a gel or detection by another method.

Figure 1

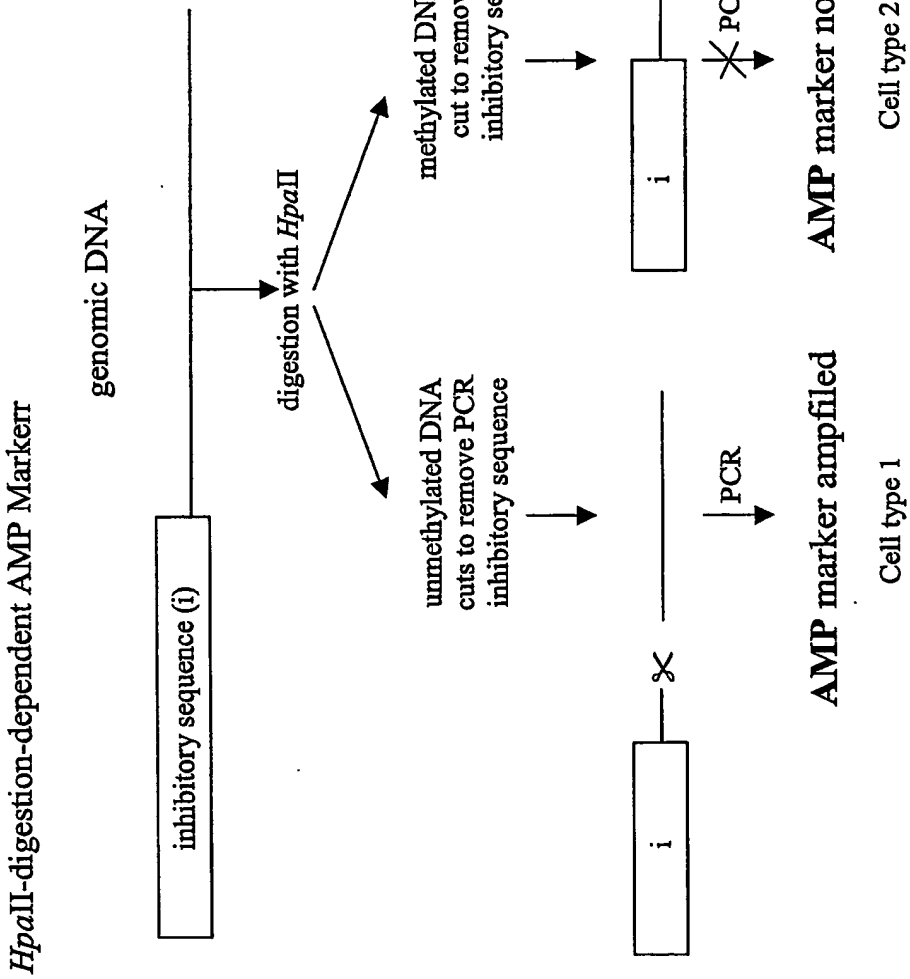

HpaII-digestion-dependent AMP Markerr
genomic DNA
inhibitory sequence (i)
digestion with HpaII
unmethylated DNA cuts to remove PCR inhibitory sequence
methylated DNA does not cut to remove PCR inhibitory sequence
PCR
PCR
i
i
AMP marker ampfiled
AMP marker not amplified
Cell type 1
Cell type 2

Figure 2

BEST AVAILABLE COPY

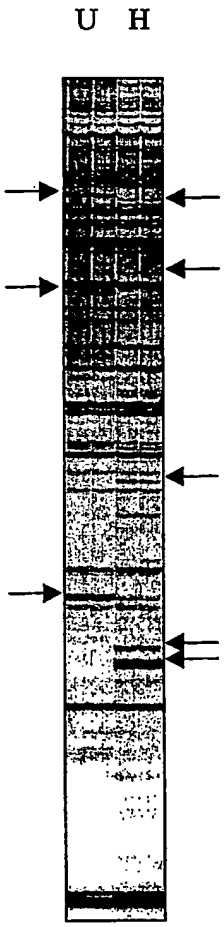


Figure 3

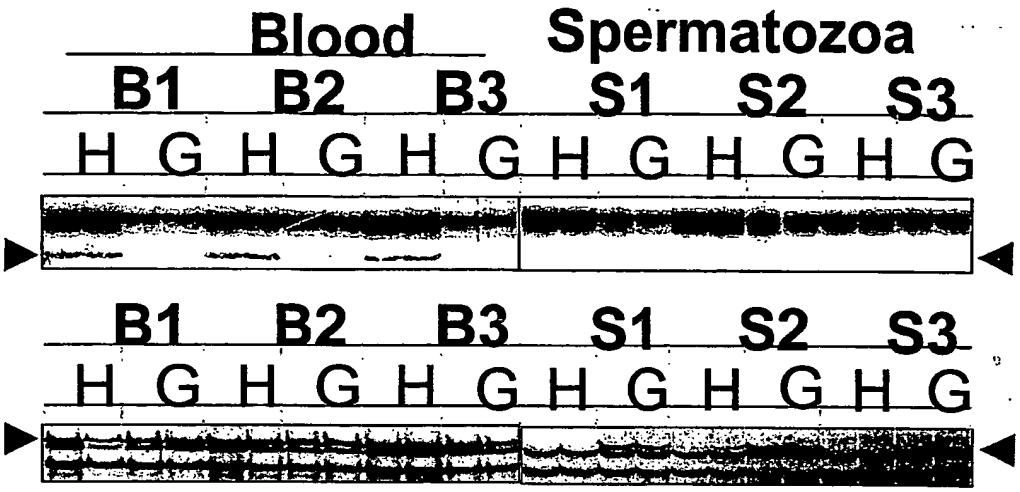
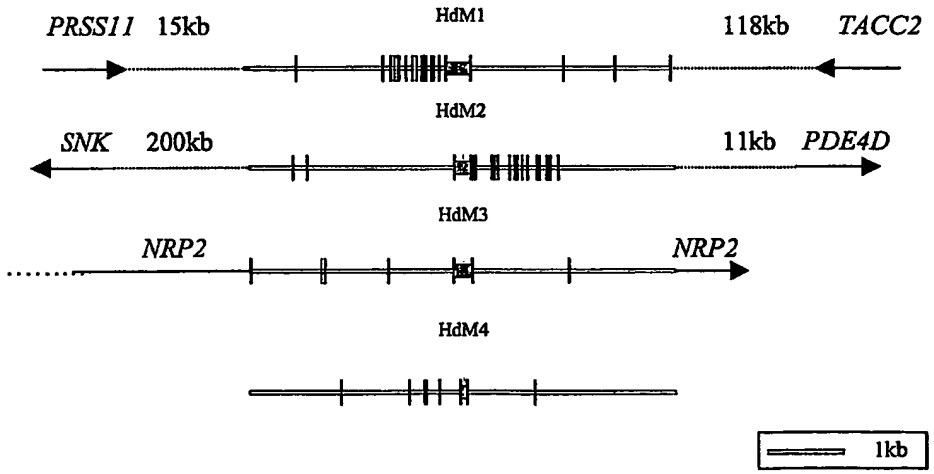


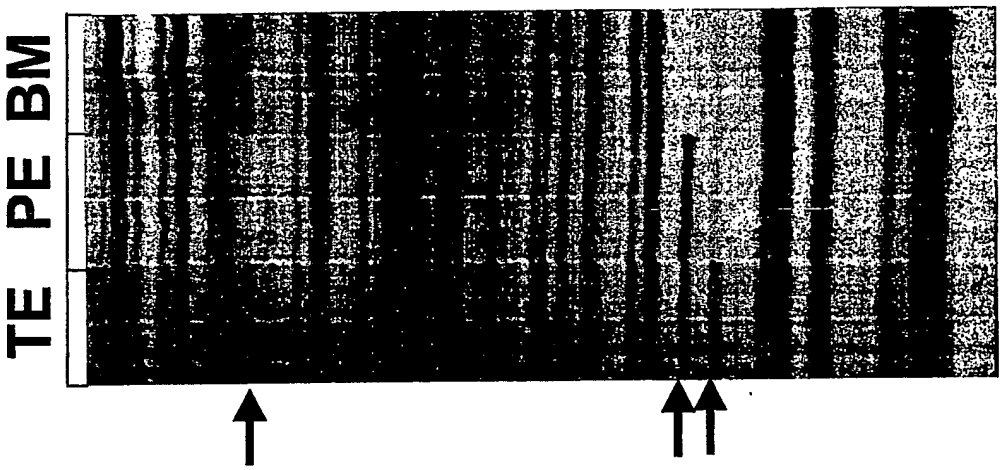
Figure 4



PRSS11 - protease, serine 11 (IGF binding protein)
TACC2 - Transforming, acidic coiled containing protein 2
SNK - Serum-inducible kinase
PDE4D - Phosphodiesterase 4D
NRP2 - Neuropilin 2

Figure 5

BEST AVAILABLE COPY

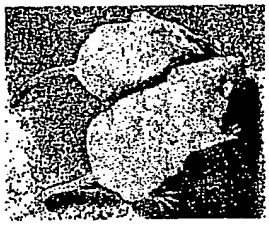


Distinct methylomes between tissue types in mouse

TE - Testis
PE - Peritoneal fat pad
BM - Bone marrow

Figure 6

BEST AVAILABLE COPY



Widespread DNA methylation polymorphisms in mouse cell lines

Genomic *Hpa*II

C5 C4 C3 C2 C1 P C5 C4 C3 C2 C1 P

M

C5 C4 C3 C2 C1 P C5 C4 C3 C2 C1 P

D

C5 C4 C3 C2 C1 P

Methylation polymorphisms:

- ~1% between the parent and a sub-cultured cell line

Figure 7

BEST AVAILABLE COPY

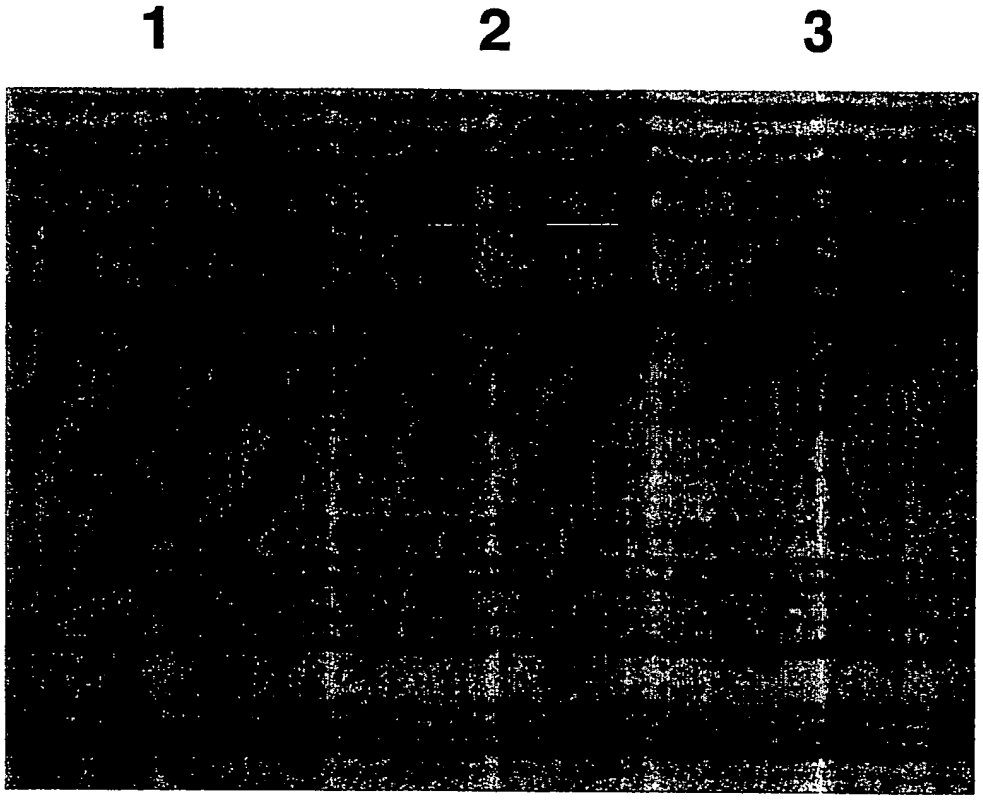


Figure 8

BEST AVAILABLE COPY

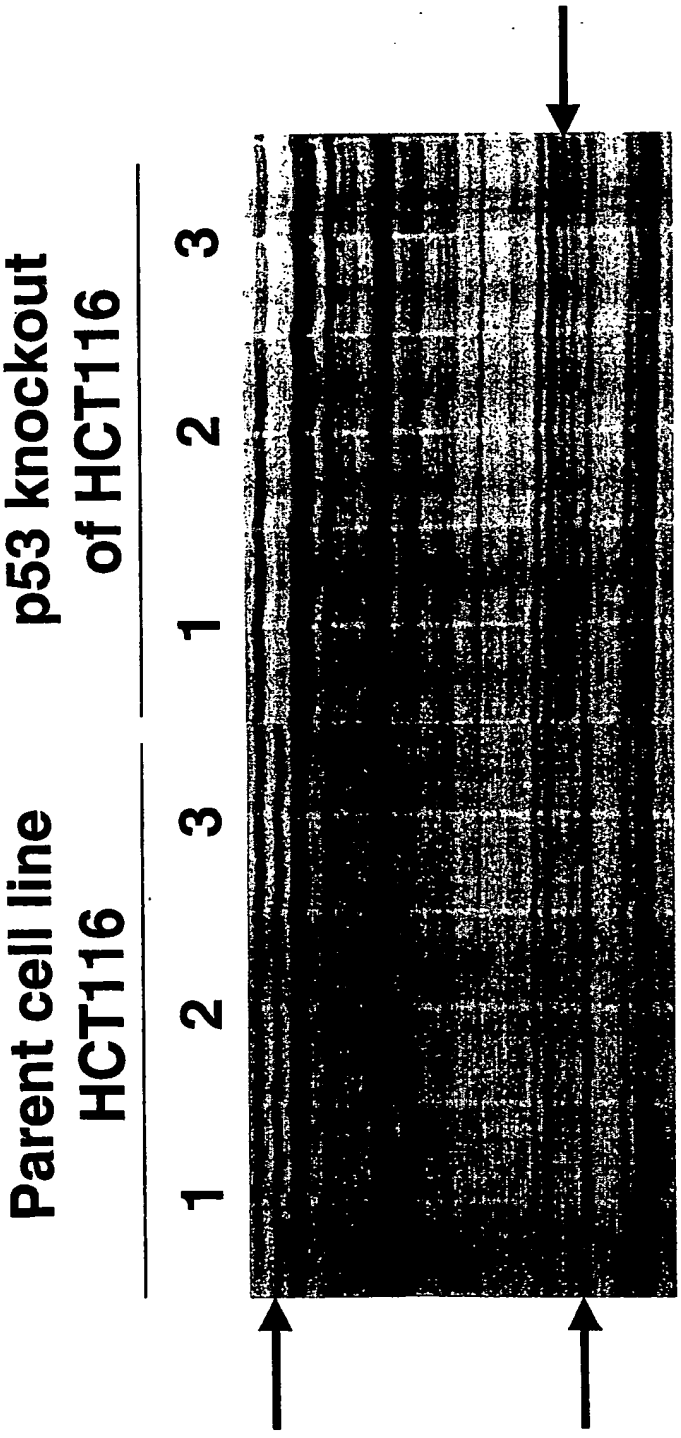


Figure 9

BEST AVAILABLE COPY

Figure 10

Southern hybridization confirms tissue-specific AMPs (Amplified Methylation Polymorphisms) between sperm (S) and blood (B)

Methylation insensitive enzymes: *Eco* RI (RI) and *Eco* RV (RV)

Methylation sensitive enzyme: *Hpa* II (Hp) 5'- CCGG-3'

Hpa II sites = vertical lines on map; * = methylated site

SEQUENCE LISTING

<110> The University of Queensland
Carroll, Bernard B (US only)
Harrison, Dion K (US only)
Aung, Hnin T (US only)

<120> A method

<130> 2564280/EJH

<140> International
<141> 2002-09-13

<150> AU PR7685
<151> 2001-09-14

<160> 38

<170> PatentIn version 3.1

<210> 1
<211> 10
<212> DNA
<213> primer

<400> 1
tttgcccgga 10

<210> 2
<211> 10
<212> DNA
<213> primer

<400> 2
tggaccggtg 10

<210> 3
<211> 10
<212> DNA
<213> primer

<400> 3
acccggtcac 10

<210> 4
<211> 10
<212> DNA
<213> primer

<400> 4
aacccgggaa 10

<210> 5
<211> 10
<212> DNA
<213> primer

<400> 5
ttcccgggtt 10

<210> 6
<211> 10
<212> DNA
<213> primer

<400> 6
tttgcccggt 10

<210> 7
<211> 10
<212> DNA
<213> primer

<400> 7
cccggcataa 10

<210> 8
<211> 10
<212> DNA
<213> primer

<400> 8
cacccggatg 10

<210> 9
<211> 10
<212> DNA
<213> primer

<400> 9
tcagtccggg 10

<210> 10
<211> 10
<212> DNA
<213> primer

<400> 10
tgccggcttg 10

<210> 11
<211> 10
<212> DNA

<213> primer

<400> 11
ccccggtaac 10

<210> 12
<211> 10
<212> DNA
<213> primer

<400> 12
cagtgccggt 10

<210> 13
<211> 10
<212> DNA
<213> primer

<400> 13
accggcttgt 10

<210> 14
<211> 10
<212> DNA
<213> primer

<400> 14
gtccggagtg 10

<210> 15
<211> 10
<212> DNA
<213> primer

<400> 15
acaccggaac 10

<210> 16
<211> 10
<212> DNA
<213> primer

<400> 16
cccggatggt 10

<210> 17
<211> 10
<212> DNA
<213> primer

<400> 17

aagaccggga 10

<210> 18
<211> 10
<212> DNA
<213> primer

<400> 18
tcccggtgag 10

<210> 19
<211> 10
<212> DNA
<213> primer

<400> 19
gaatccggca 10

<210> 20
<211> 10
<212> DNA
<213> primer

<400> 20
acccggaaac 10

<210> 21
<211> 10
<212> DNA
<213> primer

<400> 21
tgccggttca 10

<210> 22
<211> 10
<212> DNA
<213> primer

<400> 22
agccgggtaa 10

<210> 23
<211> 10
<212> DNA
<213> primer

<400> 23
cccggataca 10

<210> 24
<211> 10
<212> DNA
<213> primer

<400> 24
cccggaagag 10

<210> 25
<211> 10
<212> DNA
<213> primer

<400> 25
ctaccggcac 10

<210> 26
<211> 10
<212> DNA
<213> primer

<400> 26
acctccggtc 10

<210> 27
<211> 10
<212> DNA
<213> primer

<400> 27
ctccggatca 10

<210> 28
<211> 10
<212> DNA
<213> primer

<400> 28
tttccgggag 10

<210> 29
<211> 10
<212> DNA
<213> primer

<400> 29
aggccggtca 10

<210> 30
<211> 10
<212> DNA

<213> primer

<400> 30
caaccggtct 10

<210> 31
<211> 10
<212> DNA
<213> primer

<400> 31
ccgccggtaa 10

<210> 32
<211> 10
<212> DNA
<213> primer

<400> 32
tccgggactc 10

<210> 33
<211> 396
<212> DNA
<213> markers

<400> 33
tttcccgggt tgcctatcgc gtacttggca cacagttaag tgctcgagaa attgctaaac 60

actcaacgaa tatttaccag aatacatatt tatgctccac cgttgctgaa agaatttcga 120

ctaacccttc tttaagtctc tcttaccctc ctgtactatt aacaccattg taagtaggca 180

gtggagctac atttgcaggg agagagaaga gcgaaaaatt ggcgcgcaga gctgtaaagc 240

gtttttctcc aggtcacacc gcagcctgta aatttgtgac acttggaccg attatgggac 300

agcttgcaaa caaattatct gcaaactaac gaagcctatt tttaaagatc cctgttcctc 360

gcagctctag gcgatgacat cagagaaccc gggaaa 396

<210> 34
<211> 259
<212> DNA
<213> markers

<400> 34
taacccggga aataaattga gaggtttttt aggcttactg catgttggct attctgctcc 60

atagagggag gttcctgttt ccttgctttg cctcccctag tcacgccaga taaacatttt 120

ccaaaagcaa tttgacgcgc taaatttaag tatctcccaa gacacagggt tcctaagtaa 180

cactgagccc ttgcagcaga aacccagtag ggtccatggg cattgcatgt ttaaaggtgt 240

tcttgggatt cccgggtta 259

<210> 35
<211> 299
<212> DNA
<213> markers

<400> 35
ttggaccggt gtccagtcaa gggaataagg aggtgcattc gatgactgga gttagggcac 60

tttaaatgca ggactgaaaa tcagaaagga cgtgttctcc cccgagaaag cagagagagc 120

agagcaaagc gagaagaatt ggggatgctt gcctcatcca ggttataaaa aaaaaaaaaa 180

ggcctccagc cttgggtgat ggcaatgttg tttgtgctgt tacaatgcag cagattttac 240

attttgccct ctttgcaaac ctattagcgg acactgctga agttcccaca ccggtccaa 299

<210> 36
<211> 103
<212> DNA
<213> markers

<400> 36
tcacccggat gacatgtctg ttgacaattt taccaggtaa ttaaggaggt aatttcctac 60

acggtatagg tgaattgtcc agtctttaaa cacatccggg tga 103

<210> 37
<211> 1750
<212> DNA
<213> rpA2 promoter element

<400> 37
caggcgcccg ccaccacacc cagctaattt tttatatttt tagtagagat ggggtttcac 60

cgtgttagcc agcatggtct tgatctcctg atctcgtgat cagggaaatc tgtttaatta 120

aaacacgctt tgcgatttct gcccagtgct ctgaatgtca aagtcaagag taggaggaag 180

aagaaagaag aaagcagggt gcagtggctg acgcctgtaa tcccaacact ttgggaggcc 240

aaggtgggtg gatcacttga ggtcaggagt tcgagaccag cctgaccagc agggcgaaac 300

ccccgtctct actaaaaata caaaaattag cagggcgtgg ttgcacgtac ctgtaatccc 360

agctacgccg gaggctaagg caggaaaatc gcttgaaccc gggaggcgga ggttgcagtg 420

aggtgagatc gcaccattgc actccagcct gggcaacaag agcaaaactc agtttccgaa 480

aaaaaaaaaa aaaggaaaaa gaagagaaga aaaaaacaca cagtcctgga gcccaaaggc 540

catgtctatc ttacttaacc attgaatccc agtggccagc acagggccag gaacaaagaa 600

ggcataaaca attctttttt ttttttttt ttttttttga gtcgcagtct ctctctgcca 660

cctaggctgg agtgcagtgg cagcatcaca gctcactgca gcctggacac cccagggcca 720

ggtgatcctc ccaccttagc caccccagta gctgggaccc caggcacttg ccaccacacc 780

agactaattt ttaaaaatat tttttgagag agggtctcac tatgttgtcc aggctggtct 840

caaacttcca gcctcaagcg gtcctcctgc ctcagacccc atttgctggg tttacaggca 900

tgagccacag cacctgctaa tttttcttaa atacataaat gaacataaaa ttctaacaat 960

gcatgagtat tttgaggaag gaactgacaa aatgttccac tccctatggg aggcacgtta 1020

tatgaagaat tatgaaaaat ggtcgaaatg actggagagg ccaagcctgg atgagactgg 1080

gatggggaca ggtgcgggac gaggggcacc accctcacat ctttcacaag tctgtcatag 1140

gcaagagggc gtaggtttct cacagcccca ctggggagaa tcggcaccat tggtggcatt 1200

acacgaagag aatgtgacct cctatgtaaa agaacaagca actccacgcg gtgctgtgag 1260

gctagtgctg cgagtccctg aggtgcgcaa ttcccgcacg accgtgggtg ggaaacaccg 1320

aagccaaaac tccgctacag ccctttagat gaaggcgtcg tctgattggt gatagtttgg 1380

cgcgaacctg agcacgccga acaaaggaag tgacggcaga agtcgcgcac ttgacgaggg 1440

tgggatcaca cggcgctgcg tcgcggtagt attgttctga ttggttgatt tcttgcgata 1500

ccgctctgcc agccccttgc ttccgctagt gcggagggtt ttgcccttcg taaagatggc 1560

cgcggaggct tttggagcca actgggagcg cagtacgcgt tttctggagc atgggcagag 1620

gagacaggaa caagcgtagc atccgtgagc accgattggc tgaagcgagc accccgggag 1680

ctgactggct ccgccattcg cgggaaggcg tttgtggtgc cagagaaaag tagccagagc 1740

ggcgcagtgg 1750

<210> 38
<211> 1510
<212> DNA
<213> rbp21 promoter element

<400> 38
ctgcagagag gatgaccctt gtcaccaccg tcatgtacga ggctgcttca ccactgcctc 60

actgccacca gcgtctcccg ccgcgtgcaa tacaagaaga aacatcgaac ggtcatataa 120

ggtaagaccc actaccgatt taacctatca ttcccacaat ctaatccact tatttctctt 180

cccatgatct tatcctctca tttctcctca ctacttttgc atttgtagga aacacaatga 240

caccgtcgaa gaaagctggt ggagcaccgt agccagcaat caccaaaaca cagaggggag 300

gaggtcggca gcggccatgc ggacggcgat gagacaacgc gacgcaaaga gggaggagga 360

cgttggcgat catgctggtg ttggcggagg aggtcactgg ccatgcgaat gacagcgggg 420

cagcgcaaca caaaaagggg ggaggatgcc ggcgaccacg ctagtaccat gaagcaagat 480

gatgtgaaag ggaggaccgg acgagggttg gacctctgcc gccgacgtga agagcgtgat 540

gtgtagaagg agatgttaga ccagatgccg acgcaactta gccctgcaag tcacccgact 600

gcatatcgct gcttgccctc gtcctcatgt acacaatcag cttgcttatc tctccatact 660

tgtcgtttgt ttcccgtggc cgaaatagaa gaagacagag gtgggttttg ttggagagtt 720

ttagtggtat tgtaggccta tttgtaattt tgttgtactt tattgtatta atcaataaag 780

gtgtttcatt ctattttgac tcaatgttga atccattgat ctcttggtgt tgcactcagt 840

atgttagaat attcattccg ttgaaacaat cttggttaag ggttggaaca tttttatctg 900

ttcggtgaaa catccgtaat attttcgttg aaacaatttt tatccgacag caccgtccaa 960

caatttacac caatttggac gtgtgataca tagcagtccc caagtgaaac tgaccaccag 1020

ttgaaaggta tacaaagtga acttattcat ctaaaagacc gcagagatgg gccgtggccg 1080

tggctgcgaa acgacagcgt tcaggcccat gagccattta ttttttaaaa aaatatttca 1140

acaaaaaaga gaacggataa aatccatcga aaaaaaaaaa ctttcctacg catcctctcc 1200

tatctccatc cacggcgagc actcatccaa accgtccatc cacgcgcaca gtacacacac 1260

atagttatcg tctctccccc cgatgagtca ccacccgtgt cttcgagaaa cgcctcgccc 1320

gacaccgtac gtgcgccacc gccgcgcctg ccgcctggac acgtccggct cctctcccgc 1380

cgcgctggcc accgtccacc ggctcccgca cacgtctccc tgtctccctc cacccatgcc 1440

gtggcaatcg agctcatctc ctcgcctcct ccggcttata aatggcggcc accaccttca 1500

cctgcttgca 1510

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU02/01262

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl. [7]: C12Q 1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC (WPIDS) AND CHEMICAL ABSTRACTS

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
SEE BELOW

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
WPIDS, CA, MEDLINE, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	American Journal of Medical Genetics, 1996, vol. 64, Carrel and Willard, "An Assay for X Inactivation Based on Differential Methylation at the Fragile X Locus, FMR1", pages 27-30	1-25, 55-68, 72-87
X	Nucleic Acids Research, 1990, vol. 18, no. 3, Singer-Sam et al., "A quantitative HpaII-PCR assay to measure methylation of DNA from a small number of cells", page 687	1-25, 55-68, 72-87
X	Methods in Enzymology, 1993, vol. 255, Zuccotti et al., "Polymerase Chain Reaction for the Detection of Methylation of a Specific CpG Site in the G6pd Gene of Mouse Embryos", pages 557-567	1-25, 55-68, 72-87

[X] Further documents are listed in the continuation of Box C [ ] See patent family annex

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
31 October 2002	12 NOV 2002

Name and mailing address of the ISA/AU

AUSTRALIAN PATENT OFFICE
PO BOX 200, WODEN ACT 2606, AUSTRALIA
E-mail address: pct@ipaustralia.gov.au
Facsimile No. (02) 6285 3929

Authorized officer

Christopher Luton
Telephone No : (02) 6283 2256

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU02/01262

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Plant Molecular Biology Reporter, 1992, vol. 10, no. 4, Chang et al., "PCR Amplification Following Restriction to Detect Site-Specific DNA Methylation", pages 362-366	1-25, 55-68, 72-87
X	Cancer Research, 1997, vol. 57, Huang et al., "Identification of DNA Methylation Markers for Human Breast Carcinomas Using the Methylation-sensitive Restriction Fingerprinting Technique", pages 1030-34	1-25, 55-68, 72-87

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU02/01262

Box I **Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos :

because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos : **25-54, 69-71**

because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

See attached.

3. ☐ Claims Nos :

because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box II **Observations where unity of invention is lacking (Continuation of item 3 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU02/01262

Supplemental Box
(To be used when the space in any of Boxes I to VIII is not sufficient)

Continuation of Box No: I

Claims 25-54 and 69-71 do not define the matter for which protection is sought in terms of the technical features of the invention (Rule 6.3(a), Part B: Rules Concerning Chapter I of the Treaty).

Claim 25 purports to define a "methylation profile". A methylation profile is mere information and is therefore not limited by the technical features of the invention as described in the specification. In addition, it is noted that the subject matter of claim 25 falls within exclusion (v) of Rule 67.1 (mere presentations of information). Moreover, claim 25 is not clear to the extent that "said method" (3rd line) lacks antecedent basis.

The kit of claim 40 is not defined in such a manner as to be limited in use to the method of the invention. Claim 40 merely encompasses multiple compartments together with instructions for use. Therefore, claim 40 is not limited by the technical features of the invention as described in the specification. Moreover, claim 40 *prima facie* encompasses known materials.

The product defined by claim 69 is not limited by the technical features of the invention as described in the specification.

The system defined by claim 70 is not limited by the technical features of the invention as described in the specification. Moreover, the system defined by claim 70 encompasses a computer that differs from other computers only on the basis of the information therein. Consequently, claim 70 also falls within exclusion (v) of Rule 67.1.

The cells of claim 71 are not limited by the technical features of the invention as described in the specification. Furthermore, claim 71 may encompass known cells. The determination of a characteristic feature of an otherwise known substance cannot render that substance novel. The cells of claim 71 are not defined in such a manner as to be isolated by a method that necessarily employs the technical features of the invention.